

# Optimal Deployment of Solar Insecticidal Lamps over Constrained Locations in Mixed-Crop Farmlands

Fan Yang, Lei Shu, *Senior Member, IEEE*, Yuli Yang, *Senior Member, IEEE*, Guangjie Han, *Senior Member, IEEE*, Simon Pearson, Kailiang Li

**Abstract**—Solar Insecticidal Lamps (SILs) play a vital role in green prevention and control of pests. By embedding SILs in Wireless Sensor Networks (WSNs), we establish a novel agricultural Internet of Things (IoT), referred to as the SIL-IoTs. In practice, the deployment of SIL nodes is determined by the geographical characteristics of an actual farmland, the constraints on the locations of SIL nodes, and the radio-wave propagation in complex agricultural environment. In this paper, we mainly focus on the constrained SIL Deployment Problem (cSILDP) in a mixed-crop farmland, where the locations used to deploy SIL nodes are a limited set of candidates located on the ridges. We formulate the cSILDP in this scenario as a Connected Set Cover (CSC) problem, and propose a Hole Aware Node Deployment Method (HANDM) based on the greedy algorithm to solve the constrained optimization problem. The HANDM is a two-phase method. In the first phase, a novel deployment strategy is utilised to guarantee only a single coverage hole in each iteration, based on which a set of suboptimal locations is found for the deployment of SIL nodes. In the second phase, according to the operations of deletion and fusion, the optimal locations are obtained to meet the requirements on complete coverage and connectivity. Experimental results show that our proposed method achieves better performance than the peer algorithms, specifically in terms of deployment cost.

**Index Terms**—Solar insecticidal lamp - internet of things (SIL-IoTs), constrained SIL deployment problem (cSILDP), connected set cover (CSC) problem, approximation algorithm, wireless sensor network (WSN), path loss model, mixed-crop farmland.

Manuscript received July 22, 2020; revised Nov 29, 2020; accepted XX XX, XX. This work was supported in part by the Research Start-Up Fund for Talent Introduction of Nanjing Agricultural University under grant number 77H0603, and in part by the National Natural Science Foundation of China under Grant 62072248, and in part by the Science and Technology Innovation Project of Xuzhou in 2018 under Grant No. KC18004. Lei Shu is the corresponding author.

F. Yang is with the College of Engineering, Nanjing Agricultural University, Nanjing, 210031, China and with School of Mathematics and Statistics, Jiangsu Normal University, Xuzhou, 221116, China (email: yangfan@jsnu.edu.cn).

L. Shu is with NAU-Lincoln Joint Research Center of Intelligent Engineering, Nanjing Agricultural University, China, 210031 and with School of Engineering in College of Science, University of Lincoln, UK, INB3235, Engineering Hub, Brayford Pool, Lincoln, LN67TS, United Kingdom (e-mail: lei.shu@ieee.org).

Y. Yang is with the School of Engineering, University of Lincoln, Lincoln LN6 7TS, United Kingdom (e-mail: yyang@lincoln.ac.uk).

G. Han is with the Department of Information and Communication Systems, Hohai University, Changzhou, 213022, China and with the College of Engineering, Nanjing Agricultural University, Nanjing, 210031, China (email: hanguangjie@gmail.com).

Simon Pearson is with the Lincoln Institute for Agri-food Technology, University of Lincoln, UK, INB3235, Engineering Hub, Brayford Pool, Lincoln, LN67TS, United Kingdom (e-mail: spearson@lincoln.ac.uk).

K. Li is with the School of Artificial Intelligence, Nanjing Agricultural University, Nanjing, 210031, China (email: kailiang\_li@njau.edu.cn).

## I. INTRODUCTION

Green prevention and control of pests is one of the important technologies to protect the safety of agricultural production, the quality of agricultural products and the agro-ecological safety. As an important measure of green prevention and control, Solar Insecticidal Lamps (SILs) have been widely used in agricultural production. With SILs embedded into Wireless Sensor Networks (WSNs), the ZigBee-based SIL nodes shown in Fig. 1 are used to establish a novel agricultural Internet of Things (IoT), referred to as the SIL-IoTs. Based on the prototype presented in Fig. 1, our team has carried out several research works on SIL-IoTs, e.g., [1]–[4]. In SIL-IoTs, SIL nodes can estimate the number of killed insects by counting the discharge times [5] and transmit the pest information to users, e.g. farmers and plant protection personnel, so as to help them make reasonable decisions on the schedule and quantity of chemical control [6].

In SIL-IoTs, the deployment of SIL nodes is one of the fundamental problems, which has a direct impact on the effectiveness of routing [7]–[9] and data fusion [10], [11] operations as well as on the accuracy of anticipated coverage in several agricultural scenarios [12], e.g., mono-crop and mixed-crop farmlands. Generally, the coverage can be classified into three groups: 1) barrier coverage, 2) target coverage and 3) area coverage [13]. The barrier coverage usually deals with the detection of pests' movement across a barrier of sensors [14]–[16]. The target coverage focuses on specific points in a region of interest (RoI) [17]–[19]. The area coverage often concerns monitoring the whole RoI [20]–[22]. Since most applications requires full coverage, the majority of current studies focus on the area coverage problem. For specific applications where the coverage is not a mandatory requirement, the other metrics can be used to evaluate the node deployment, e.g., deployment cost (the number of used nodes), connectivity, network reliability, network lifetime and network fault tolerance.

### A. Related works

In the past few decades, a series of researches have been carried out on the area coverage. A top-down survey on the area coverage in terms of coverage and connectivity has been presented in [23]–[27], and we refer the reader to these comprehensive survey papers for more detailed discussions. In this section, we provide an overview on some related works in terms of area coverage.

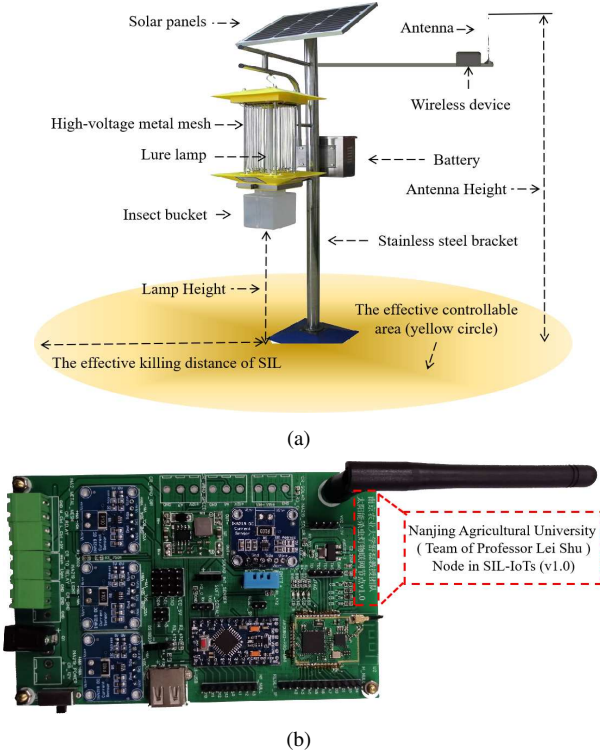


Fig. 1. A ZigBee-based SIL node. (a) The prototype of a SIL node designed with the wireless communication protocol ZigBee. (b) The wireless device in a SIL node.

### 1) Deterministic Algorithms

Kershner [28] has proven the optimal deployment pattern in an infinite plane based on the disk coverage model. Wang et al. [29] have proven the optimal deployment pattern in long-belt scenarios based on the Confident Information Coverage (CIC) model [30]. On the basis of these works, a series of node deployment methods have been proposed [31]–[35]. In [33], the Layered Deployment Model (LDM) was proposed to address the area coverage problem in a scenario with characteristics of full coverage and irregular RoI, where the centroid of irregular area was first sought as a fusion node, and then the remaining nodes were deployed in the layered form to completely cover the area. In [34], a Growth Rings Like Deployment (GRLD) scheme was proposed to deploy nodes in a bounded area. In [35], the area coverage problem was investigated in an arbitrary region with an irregular shape and inner obstacles, where a projection-based approach was proposed to minimize the number of sensor nodes while guaranteeing full coverage of such an area.

### 2) Heuristics Algorithms

Heuristics algorithms provide a better way to solve the area coverage problem, because they are able to find a near optimal solution within a reasonable time at limited calculation cost [20]–[22], [36]–[39]. In [20], according to the natural partition structure of an actual farmland, two deployment methods were proposed based on the Genetic Algorithm (GA) to solve the SIL Deployment Problem (SILDP). In [21], an improved nature-inspired GA was proposed to maximize the network lifetime and coverage rate while guaranteeing the minimum

deployment cost. In [22], the area coverage problem was formulated as the Mixed Integer Linear Program (MILP) problem, and an iterative method based on the GA was proposed to solve this problem. In [36], the normalization method and an improved GA were proposed to solve the maximum coverage sensor deployment problem in heterogeneous WSNs. In [37], an Improved Flower Pollination Algorithm (IFPA) was presented to overcome the shortcomings of slow convergence and low accuracy of the FPA algorithm. In [38], a bipopulation-based evolutionary full area coverage algorithm was proposed to achieve full coverage at the minimum deployment cost in the RoI with non-penetrable obstacles. In [39], a harmony search-based deployment algorithm was proposed to locate the optimal number of nodes, where the optimal locations are obtained to achieve the maximum coverage rate at the minimum deployment cost.

### 3) Approximation Algorithm

For a continuous monitoring area, the candidate locations used to deploy sensor nodes are actually infinite. Therefore, the continuous monitoring area is often divided into consecutive and countable grid points or cells. A sensor node can only be deployed at a grid point or the centre of a grid cell, and full coverage is achieved for all grid points being covered. The approximation algorithm is an effective method to deal with the grid-based node deployment problem [40]–[48]. In [40], a new node deployment problem was presented for the minimum energy harvesting node placement with respect to the energy neutral coverage and connectivity, which was formulated as the MILP. Later on, two deployment algorithms have been proposed to solve the presented deployment problem. In [41], the Greedy Approximation Algorithm (GAA) was proposed to address the minimal cover set problem. In [42], a Greedy-based Partial Coverage (GPC) algorithm was proposed, which utilizes the neighbour nodes to preserve the connectivity of the selected nodes and adopts the overlap between nodes to achieve the required coverage rate. In [43], the Candidate Location Based Greedy Algorithm (CLBGA) was presented to solve the full area coverage problem based on the CIC model. In [44], the Local Search (LS) algorithm was proposed to deploy the minimum number of nodes while guaranteeing full coverage and connectivity. In [45], a Greedy Placement based on Iterative Delaunay Triangulation (GPIDT) was proposed to achieve complete coverage based on the CIC model in an area with obstacles.

To compare the aforementioned works, a summary of typical studies on the area coverage is listed in Table I.

### B. Limitations and Challenges

Although the previous researches can effectively solve the node deployment problem in industrial IoTs, they have to face the following challenges when applied in agricultural IoTs, especially in SIL-IoTs.

- 1) Almost all existing methods study the node deployment problems without setting any constraints on the geographic locations of nodes, i.e., they assume that the nodes can be deployed anywhere. In fact, there are some physical constraints on the placement of nodes. For instance, in

TABLE I  
COMPARATIVE ANALYSIS OF TYPICAL STUDIES ON AREA COVERAGE.

Classification	Year	Ref	Main goals	RoI	Obstacle	Deployment Environment	Node Position Constraint
Deterministic Algorithms	2011	[33]	full area coverage and connectivity	Irregular	Yes	Homogeneous	No
	2013	[34]	full area coverage and connectivity	Irregular	No	Homogeneous	No
	2014	[35]	full area coverage and connectivity	Irregular	Yes	Homogeneous	No
Heuristics Algorithms	2018	[22]	full area coverage and connectivity	Regular	Yes	Homogeneous	No
	2019	[21]	confidential-information coverage network lifetime maximisation	Regular	Yes	Homogeneous	No
	2019	[37]	coverage maximisation full connectivity	Regular	Yes	Heterogeneous	No
	2020	[20]	full area coverage and connectivity	Irregular	Yes	Homogeneous	No
Approximation Algorithms	2015	[43]	confidential-information coverage	Regular	No	Homogeneous	No
	2015	[45]	full area coverage and connectivity	Regular	No	Homogeneous	No
	2017	[40]	energy neutral coverage and connectivity	Regular	No	Homogeneous	No
	2017	[42]	coverage maximisation	Irregular	No	Heterogeneous	No
	2021	our proposed	full area coverage and connectivity	Irregular	Yes	Heterogeneous	Yes

a WSN application for monitoring crop growth, there is a lower bound on the distance between a node and the farmland edge to avoid the edge effect on the accuracy of monitoring information [49]. As to the placement of SIL nodes, they are recommended to be deployed on the ridges for minimizing the daily maintenance cost of SILs and reducing the impact on agricultural machinery services. In general, the node deployment has to be constrained in practical agricultural applications.

- 2) Most of the previous works are based on the assumption that the node deployment environment is homogeneous with the same path loss for all locations. However, the real-world deployment environment is more complex and heterogeneous deployment environment is common in practical applications, especially in agricultural applications. Since the agricultural environment is dynamic, the path loss in the presence of dense plants is varying as crops grow. The propagation of radio waves is seriously affected by the reflection and refraction in the channels, the shadow fading due to the antenna setting, the height of crop canopy and the surrounding [50]. Therefore, the propagation characteristics of wireless channels in an actual farmland have to be considered in the SIL node deployment.
- 3) For the area coverage problem where nodes can only be deployed at a limited set of predetermined candidate locations, e.g., the grid-based area coverage problem, the use of Linear Program (LP) based deployment algorithms, although capable of finding optimal solutions [40], [44], is not efficient because the time in finding optimal solutions increases dramatically with the network size. For heuristics algorithms, e.g., [22], [37], [39], the encoding scheme and genetic operators are not suitable for the fast convergence with a small number of constraints, since the area coverage problem is a high-dimensional combinatorial optimization

problem. Moreover, for the RoI with an irregular shape, it is difficult to cover some special points including the convex vertices on the boundary of RoI, by randomly generated solutions within a limited number of iterations. For approximation algorithms, many equivalent candidates are available in the solution space, an inappropriate choice will increase the dimension of solution vector. Additionally, due to more coverage holes generated, more nodes are needed to achieve full coverage, as shown in Fig. 2. However, most of previous works, e.g., [41]–[45], have not considered the impact of new nodes added in each iteration on the number of coverage holes, which increases the dimension of solution vector.

### C. Contributions and Characteristics

In our previous work [20], the SILDP was studied, where the constraints on the SIL node locations and the heterogeneous deployment environment have not been considered. However, the SILDP subject to the requirements on the locations of SIL nodes and the upper bound on the distance of internode communications in an actual farmland is intrinsically harder than its unconstrained counterparts. As the first step toward addressing this challenging deployment problem, the constrained SIL Deployment Problem (cSILDP) is studied in this paper, where the placement of SIL nodes is restricted to a set of candidates located on the ridges<sup>1</sup>. An actual farmland shown in Fig. 3 is studied in this paper, and the main contributions of this paper are summarized as follows:

<sup>1</sup>Although some researches on the relay node placement have set constraints on the geographic locations, e.g., [51]–[57], these studies focus on either the network connectivity or the network lifetime, i.e., how to place the minimum number of relay nodes within a WSN for meeting the requirements on connectivity or survivability. Since the cSILDP focuses on the coverage, these studies cannot be applied to solve the cSILDP.

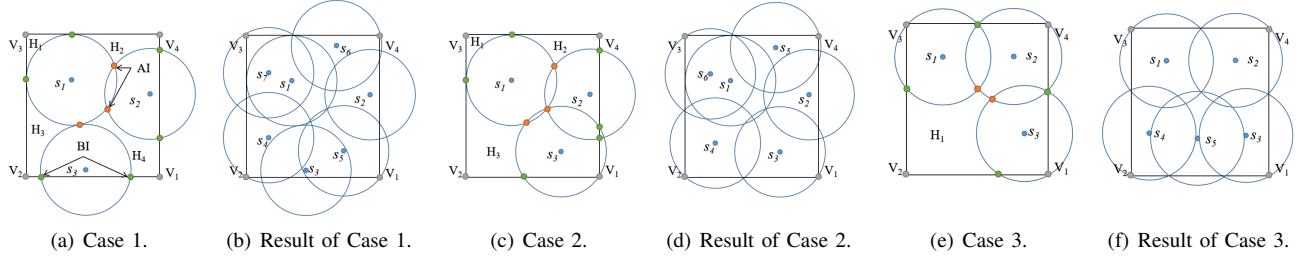


Fig. 2. An instance of coverage holes, with two deployed nodes  $s_1$  and  $s_2$ . Blue dots denote deployed nodes and a blue circle indicates the sensing range. Green dots denote Boundary Intersection (BI) between circle and boundary. Orange dots denote Arc Intersection (AI) between a pair of circles. Gray dots denote the vertex of RoI. In (a),  $s_3$  is selected to be deployed since it has the maximum coverage rate, where four coverage holes,  $H_1, H_2, H_3, H_4$ , are generated. Following the deployment strategy, four additional nodes are needed to fully cover the whole square, as shown in (b). In (c),  $s_3$  is not the optimal one in terms of coverage rate, but there are three coverage holes, i.e.,  $H_1, H_2, H_3$ , where three additional nodes are needed to completely cover the holes, as shown in (d). In (e), only one coverage hole is generated after adding  $s_3$  to the solution, where five nodes are needed to achieve the full coverage, as shown in (f). Therefore, the more coverage holes are generated, the more nodes are needed to completely cover these holes.

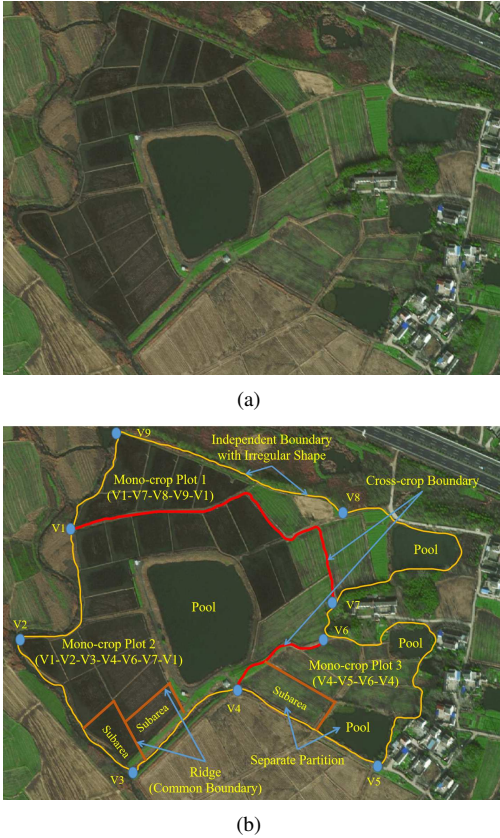


Fig. 3. An actual farmland studied in this paper. (a) An overview of the farmland with characteristics of mixed crop, irregular boundary, partition structure and penetrable obstacles (e.g., water pools). This farmland is located in Babaqiao Town, Nanjing, China.  $118^{\circ}52'E, 32^{\circ}24'N$ . (b) The distribution of mono-crop plots in the farmland. According to the partition structure caused by natural physiognomy feature, the farmland can be divided into subareas by ridges. Each subarea is a separate partition and the mono-crop plots include a set of separate partitions with same crop.

- 1) By using the candidate locations, we simultaneously approximate the constraints enforced by special requirements on the locations of SIL nodes and the internode distance bound for insect killing or communications. In other words, by adopting a discrete optimization problem, a non-convex continuous optimization problem is approximated.
- 2) According to the previous works on the analysis of prop-

agation characteristics of wireless channels in different mono-crop plots, e.g., wheat [58], corn [59], rape [60] and rice [61], we utilize a physical layer model of point-to-point communications to derive the maximum transmission distance between SIL nodes for a certain crop in maturity stage<sup>2</sup>.

- 3) Since the phototaxis of pests is different not only in species, but also in geography and seasons, various kinds of pests have different requirements on the effective killing distance<sup>3</sup>. In addition, the maximum transmission distance of a SIL node is various in different mono-crop plots. Therefore, the cSILDP in a mixed-crop farmland can be regarded as the heterogeneous node deployment problem. We formulate the cSILDP as a Connected Set Cover (CSC) problem and prove that it is NP-hard. To the best of our knowledge, this paper are the first to prove that the cSILDP is NP-hard.
- 4) We propose a Hole Aware Node Deployment Method (HANDM) based on the greedy algorithm to solve the cSILDP in a mixed-crop farmland. The HANDM is a two-phase method. In the first phase, we put forth a novel deployment strategy to guarantee only a single coverage hole in each iteration, based on which we deploy a set of SIL nodes at suboptimal locations. In the second phase, considering the existence of redundant nodes and the pair of SIL nodes with high overlap rate, we design two operations, i.e., deletion and fusion operation, to finetune the deployed SIL nodes toward the optimal locations for reducing the deployment cost while meeting the requirements on full coverage and connectivity.

#### D. Organization

The rest of this paper is organized as follows. In Section II, we introduce the models and definitions used in the design of SIL-IoTs. In Section III, we formulate the cSILDP. Next,

<sup>2</sup>The reason for choosing maturity stage as a representative of the worst-case scenario is that the plant stem and foliage are fully developed and the propagation environments are the worst. Therefore, if a pair of SIL nodes can successfully communicate in maturity stage, they can certainly communicate during the whole growth cycle.

<sup>3</sup>The effective killing distance means that the number of a certain kind of pests killed by a SIL accounts for 61.8% of the total number of pests killed by the SIL, within this distance [62].

the proposed deployment method is described in Section IV. We provide simulation results and discussions in Section V. Finally, we conclude this paper in Section VI.

## II. PRELIMINARIES

### A. Network Model

Consider there are  $((m+1)/\zeta) \times ((n+1)/\zeta)$  grid points in an  $m \times n$  rectangle network model, where  $m$  and  $n$  are the length and width of the network, respectively. The uniform length of each grid is  $\zeta$ , which is determined by users' requirement on the accuracy. If a higher precision is needed,  $\zeta$  is set to a smaller value at the cost of higher calculation complexity. Otherwise, a larger value is set to achieve lower calculation complexity.

There is a farmland  $A$  made up of three mono-crop plots in this rectangle network model, as show in Fig. 4, where  $A = \{A_f, A_{if}\}$  with  $A_f = \{A_f^1, A_f^2, \dots, A_f^{|A_f|}\}$  and  $A_{if} = \{A_{if}^1, A_{if}^2, \dots, A_{if}^{|A_{if}|}\}$ . Moreover,  $A_c^1 \cup A_c^2 \cup A_c^3 = A_f$ . Herein,  $A_f^i$  represents the  $i^{\text{th}}$  feasible subarea, consisting of the ridge  $A_R^i$  and the planting region  $A_P^i$ .  $A_{if}^j$  represents the  $j^{\text{th}}$  infeasible subarea (e.g., water pool), consisting of the ridge  $A_R^j$  and the obstacle  $A_O^j$ . As shown in Fig. 4,  $A_R$  is an area where SILs can be deployed, and  $A_P$  and  $A_O$  are the areas where SILs cannot be deployed, but wireless signals can pass through.  $A_c^k$  represents the  $k^{\text{th}}$  mono-crop plot, which is a subset of  $A_f$ , i.e.,  $A_c^k = \{A_f^1, A_f^2, \dots, A_f^{|A_c^k|}\}$ . For any two different mono-crop plots  $A_c^i$  and  $A_c^j$ ,  $A_c^i \cap A_c^j = \emptyset$ .

Notations used in this paper are listed in Table II.

### B. Definitions and Functions

**Definition 1.** A mono-crop plot is a set of feasible subareas, which is a homogeneous unit that can be described using the same channel propagation model and effective killing distance.

**Definition 2.** Candidate locations in the same mono-crop plot have the same features of effective killing distance and communication distance. Otherwise, they do not have the same features.

**Definition 3.** Cross-crop boundary is a common boundary between adjacent mono-crop plots;

**Definition 4.** Point of Interest (PoI) is a grid point in feasible subareas, which is the intended and monitored points. Let  $U$  be a set of PoIs.

**Definition 5.** Full coverage of  $A$  represents that all PoIs are covered by a set of SILs, denoted by  $S = \{s_1, s_2, \dots, s_{|S|}\}$ , with  $s_i$  standing for the  $i^{\text{th}}$  SIL node in  $S$ .

**Definition 6.** Full connectivity of a network means any two SIL nodes in the network can communicate with each other in the single-hop or multi-hop manner.

**Definition 7.** A SIL node  $s_i$  is a redundant one, if its effective controllable area shown in Fig. 1 is overlapped by other SIL nodes.

**Definition 8.** A PoI is defined as a Boundary Point (BP), if one of its one-hop neighbour grid points is not in  $A$ . Let  $B$  be a

TABLE II  
NOTATIONS

Notation	Description
$m/n$	Length / width of the network
$\zeta$	Uniform length of grid
$A_f / A_{if}$	Feasible / Infeasible subarea
$A_P / A_R / A_O$	Planting region / Ridge area / Obstacle area
$U/B$	Set of PoIs / BPs
$V$	Boundary coverage vector (BCV)
$A_c$	Mono-crop plot
$A_A / A_A^i$	Size of area $A$ / $A_f^i$
$S$	Set of SILs (Solution)
$G$	Adjacency matrix of graph induced by $S$
$CP$	Set of vertices of the polygon shaping the actual farmland
$L$	Set of candidate locations
$L'$	Set of candidate locations that cover one point in $I_E$
$L''$	Set of candidate locations that satisfy the full coverage requirement
$L'''$	Set of candidate locations that satisfy the full connectivity requirement
$L_{CP}$	Set of candidate locations that cover any point in $CP$
$L_{JP}$	Set of candidate locations that generate only one coverage hole
$L_{Com}$	Set of candidate locations that can communicate with any SIL node in $S$
$\rho$	Candidate location density
$R$	Effective killing distance
$R_c$	Maximum transmission distance
$P_i$	Transmit power at SIL $s_i$
$\theta_{i,j}$	The channel coefficient between the transmitter $s_i$ and the receiver $s_j$
$\phi_{i,j}$	Instantaneous received signal-to-noise ratio at SIL $s_j$ with mean value $\bar{\phi}_{i,j}$
$n(t)$	A zero-mean additive white Gaussian noise (AWGN) with variance $\sigma^2$
$\gamma_{th}$	Predefined receiver sensitivity threshold
$\alpha$	Path loss exponent
$N$	Deployment cost (Size of cover set / Dimension of solution)
$P_{cov}^{A_f}(S, j)$	Coverage probability of PoI $j$ in $A_f$
$P_{ove}^{A_f}(S, j)$	Overlap probability of PoI $j$ in $A_f$
$C_{cov}^{A_f}(S)$	Coverage rate of $A$
$C_{ove}^{A_f}(S)$	Overlap rate of $A$
$p_{size}$	Size of initial population
$d(i, j)$	Euclidean distance between $i$ and $j$
$ \cdot $	Number of cells in a set, i.e., cardinality
$\mathbb{E}(\cdot)$	Expectation operator

set of BPs, and BPs in  $B$  are in clockwise or counterclockwise order.

**Definition 9.** Given a solution  $S$  and a  $[0,1]$  row vector whose dimension is  $|B|$ , its  $m^{\text{th}}$  element is 1 only if the  $m^{\text{th}}$  BP in  $B$  can be covered by any SIL node in  $S$ . Otherwise, this element is 0. This row vector is defined as the Boundary Coverage Vector (BCV). Let  $V$  be the BCV.

**Definition 10.** Given a BCV  $V$ , if the values of adjacent elements (the first and last elements are also considered as the adjacent elements) in  $V$  are different, Jump Point (JP) is the index of these two elements whose value is 0.

**Definition 11.** A BP is defined as a Boundary Intersection (BI), if its index in  $B$  is equal to JP. A BI is an Effective



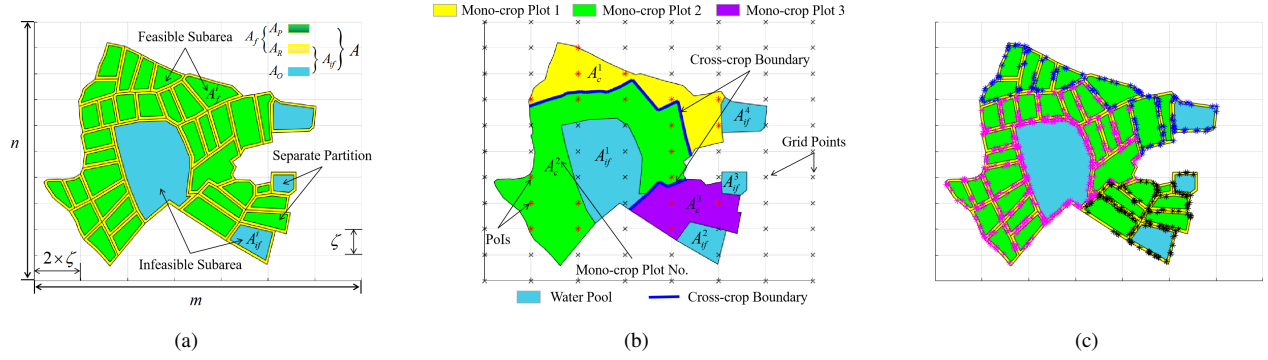


Fig. 4. Map 1: (a) the map used to model the actual farmland in Fig. 3, (b) the distribution of mono-crop plots, (c) the candidate locations where SIL nodes can only be deployed, where blue stars, pink stars and black stars represent the candidate locations in  $A_c^1$ ,  $A_c^2$  and  $A_c^3$ , respectively.

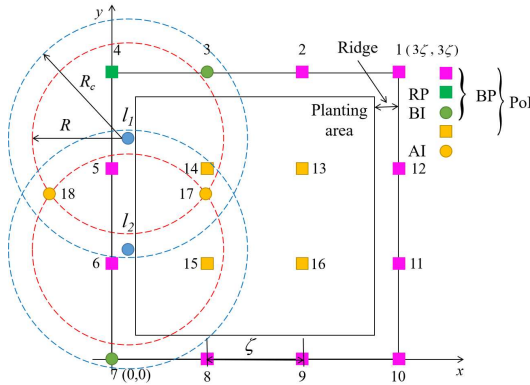


Fig. 5. An instance of deployment map with  $U = \{1, 2, \dots, 15, 16\}$ ,  $|U| = 16$ .  $B = \{1, 2, \dots, 11, 12\}$ ,  $|B| = 12$ , where the BPs are marked counterclockwise. The blue dot represents the candidate location. The red dash circle represents the effective killing distance. The blue dash circle represents the maximum communication distance.

*BI (EBI), if it cannot be covered by any SIL node in  $S$ . The Arc Intersection (AI) is an intersection of two sensing circles<sup>4</sup>. An AI is defined as an Effective AI (EAI), if it is within the polygon<sup>5</sup> and the distance from it to any deployed SIL node is not less than the node's effective killing distance. Let  $I_A$  be a set of AIs,  $I_B$  be a set of BIs,  $I_{EA}$  be a set of EAIs,  $I_{EB}$  be a set of EBIs,  $I$  be a set of intersection points,  $I_E$  be a set of effective intersection points, i.e.,  $I = I_A \cup I_B$ ,  $I_{EA} \subseteq I_A$ ,  $I_{EB} \subseteq I_B$ ,  $I_E \subseteq I$ ,  $I_E = I_{EA} \cup I_{EB}$ .*

**Definition 12.** A given convex point of a polygon is defined as the Reference point (RP), if this point is on the boundary of minimum bounding rectangle (MBR) of the polygon and can be covered by the first SIL node. The boundary where an RP is located is defined as the Reference Boundary (RB).

**Definition 13.** Coverage hole is an area that is uncovered, but enclosed by sensing circles and the boundary of polygon.

<sup>4</sup>For convenience, sensing circle is used to represent a SIL's effective controllable area. The radius of sensing circle is equal to the effective killing distance.

<sup>5</sup>As the shape of an actual farmland can be approximated by a polygon, which is used to represent the shape of an actual farmland in the remaining of this paper. Therefore, for a point located in the farmland, we say it is within the polygon.

TABLE III  
FUNCTIONS

Function	Description
$X(l)$	Return the set of PoIs within the effective killing distance of candidate location $l$
$Y(p)$	Return the set of candidate locations that are able to cover point $p$
$Z(l)$	Return the set of candidate locations that can communicate with candidate location $l$
$W(S)$	Return the BCV $V$ that is calculated by $S$
$J(V)$	Return the set of JPs in $V$
$D(i)$	Return the distance between the effective intersection point $i$ and RB

However, the uncovered areas that are enclosed by many sensing circles are not the so-called coverage holes in this paper.

Let  $X(l)$  represent the set of PoIs within the effective killing distance of candidate location  $l$ . Conversely,  $Y(p)$  represents the set of candidate locations that are able to cover point  $p$ , and  $Z(l)$  represents the set of candidate locations that can communicate with candidate location  $l$ .  $W(S)$  represents a BCV  $V$  that is calculated by  $S$ , i.e.,  $V = W(S)$ , and  $J(V)$  represents the set of JPs in  $V$ . The functions used in this paper are listed in Table III.

For example, in Fig. 5,  $S = \{l_1, l_2\}$ ,  $U = \{1, 2, \dots, 15, 16\}$ ,  $B = \{1, 2, \dots, 11, 12\}$  where  $U(1) = B(1) = (3\zeta, 3\zeta)$ ,  $|U| = 16$ ,  $|B| = 12$ , and the BPs are marked counter-clockwise. The PoIs that can be covered by  $l_1$  are  $X(l_1) = \{4, 5, 14\}$ . The candidate locations that can cover point 17 are  $Y(17) = \{l_1, l_2\}$ . The candidate location that can communicate with  $l_2$  is  $Z(l_2) = \{l_1\}$ . Moreover,  $W(S) = \{0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0\}$ , and  $J(V) = \{3, 7\}$ , where  $V = W(S)$ ,  $|J(V)| = 2$ ,  $I_{EB} = \{3, 7\}$ ,  $I_A = \{17, 18\}$ ,  $I_E = \{3, 7, 17\}$ .

**Lemma 1.** *Given a solution  $S$  and a BCV  $V$ , if  $|J(V)| \leq 2$  where  $V = W(S)$ , the network built by  $S$  will have only one coverage hole.*

*Proof.* See Appendix A.

### C. Effective Killing Distance

In our survey paper [6], we have found that various kinds of pests have different requirements on the effective killing distance [62]–[65]. Therefore, in this paper, we assume that for a certain kind of pests, the SIL node has a fixed effective killing distance. Moreover, we also assume that any point  $j$  within the effective killing distance of SIL  $s_i$  can be directly covered by  $s_i$ . Thus, the probability that Point  $j$  in  $A_f$  is covered by  $s_i$  can be derived as

$$f(s_i, j) = \begin{cases} 1, & d(s_i, j) \leq R^{s_i} \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where  $R^{s_i}$  is the effective killing distance of SIL node  $s_i$ , and the operation of  $d(\cdot)$  is used to calculate the Euclidean distance between  $s_i$  and  $j$ .

### D. Communication Model

In this part, the point-to-point communication model between a pair of SIL nodes in the network is presented. Let  $s_i$  and  $s_j$  be the transmitter SIL node and receiver SIL node, respectively. The channel coefficient between these two nodes is expressed as  $\theta_{i,j}$ . Thus, the received signal at node  $s_j$  is defined as [66]

$$r_j(t) = \sqrt{P_i} \theta_{i,j} s_i(t) + n(t), \quad (2)$$

where  $P_i$  is the transmit power at node  $s_i$ ,  $s_i(t)$  is the symbol transmitted by  $s_i$  at time  $t$ , and  $n(t)$  is the AWGN with zero mean and variance  $\sigma^2$ . The value of  $\sigma^2$  is generally normalized to 1 [67].

Let  $\phi_{i,j}$  and  $\bar{\phi}_{i,j}$  be the instantaneous received signal-to-noise ratio (SNR) and the average received SNR at SIL  $s_j$ , which can be expressed as [67]

$$\phi_{i,j} = \frac{P_i \theta_{i,j}^2}{\sigma^2} \quad (3)$$

and

$$\bar{\phi}_{i,j} = \frac{P_i \mathbb{E}(\theta_{i,j}^2)}{\sigma^2} \quad (4)$$

respectively, where  $\mathbb{E}(\cdot)$  denotes the expectation operator and thus  $\mathbb{E}(\theta_{i,j}^2)$  is the variance of channel coefficient. According to the distance dependent path loss model, the variance of channel coefficient can be expressed as [68]

$$\mathbb{E}(\theta_{i,j}^2) = \left( \frac{d_0}{d(s_i, s_j)} \right)^\alpha, \quad (5)$$

where  $d_0$  and  $\alpha$  represent the reference distance and the path loss exponent, respectively. The value of  $d_0$  is usually set to 1 [68]. Therefore, (4) can be rewritten as

$$\bar{\phi}_{i,j} = \frac{P_i}{\sigma^2} \left( \frac{d_0}{d(s_i, s_j)} \right)^\alpha. \quad (6)$$

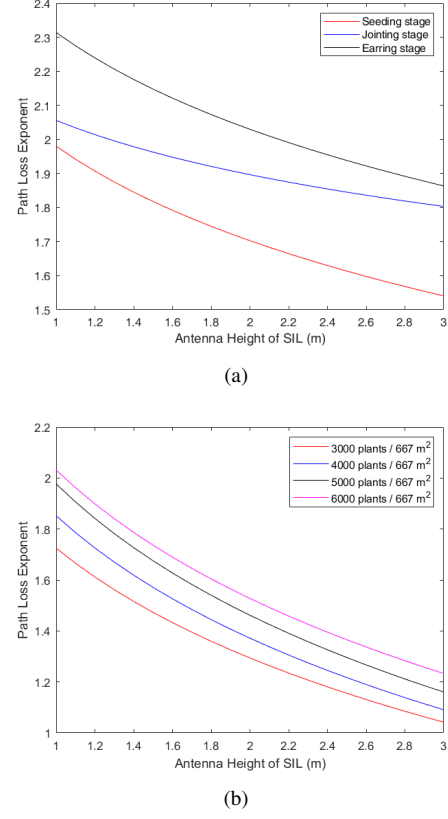


Fig. 6. (a) Impact of crop growth stages on  $\alpha$  [58]. (b) Impact of plant density on  $\alpha$  [69].

### E. Path Loss Model

Considering the effects of frequency and propagation distance on wireless channels, the presence of plants between a pair of transmitter and receiver, together with the reflection, refraction, diffraction and absorption of signals by plants, the log-linear model is often adopted to characterise the signal transmissions in agricultural environment [12], [59], [60]. The path loss based on the log-linear model is formulated as

$$P_L(d) = P_L(d_0) + 10\alpha \log_{10}(d(s_i, s_j)) + Z_{\sigma^2}, \quad (7)$$

where  $P_L(d)$  is path loss in dB, and  $P_L(d_0)$  is the reference distance path loss in dB, which is related to the specific application environment. Besides,  $\alpha$  is the path loss exponent, and  $Z_{\sigma^2}$  is the Gaussian random variable with variance of  $\sigma^2$  caused by shadow fading.

As an important parameter affecting the path loss,  $\alpha$  is highly related with the crop growth stage [58], canopy height [61], plant density [69] and surrounding environment [70], e.g., shadow fading caused by antenna height, as shown in Fig. 6. In this paper, we use various  $\alpha$  to represent the path loss in different mono-crop plots.

### F. Maximum Transmission Distance

As is known, if two nodes are unable to communicate with each other in the free space, they will certainly fail to communicate in a practical agricultural environment. Since  $P$  and  $\alpha$  are two important parameters to judge the possibility of

communication, as shown in (6), the maximum transmission distance  $R_c^{s_i}$  of SIL node  $s_i$  is determined by these two parameters.

For a SIL node  $s_j$ , it can successfully receive the signals sent from  $s_i$  in a real-world agricultural environment should meet the distance requirement, i.e.,  $d(s_i, s_j) \leq R_c^{s_i}$ . The following lemma shows how to determine the value of  $R_c^{s_i}$ .

**Lemma 2.** *If  $s_j$  is able to successfully receive the signals sent from  $s_i$  in a real-world agricultural environment, the distance from  $s_i$  to  $s_j$  must be less than  $R_c^{s_i}$  given by*

$$R_c^{s_i} = d_0 \alpha^i \sqrt{\frac{P_i}{\sigma^2 \gamma_{th}}}, \quad (8)$$

where  $\gamma_{th}$  is a predefined receiver sensitivity threshold and  $\alpha^i$  is the path loss exponent in the mono-crop plot for  $s_i$ .

*Proof.* See Appendix B.  $\square$

Concerning the asymmetric link between  $s_i$  and  $s_j$ , these two nodes can communicate with each other only if

$$d(s_i, s_j) \leq \min(R_c^{s_i}, R_c^{s_j}). \quad (9)$$

### G. Metrics of SIL Deployment

In this paper, three metrics are considered for the SIL node deployment, including coverage, overlap and connectivity.

#### 1) Coverage and Overlap

Similar with [20], the grid-based computing method is adopted to calculate the coverage and overlap. Let  $P_{cov}^{A_f}(S, k)$  be the probability of  $k$  covered by any SIL node in  $S$ , and  $P_{ove}^{A_f}(S, k)$  be the probability of  $k$  covered by at least two SIL nodes in  $S$ , where  $k$  is a PoI in  $A_f$ , these two probabilities can be expressed as

$$P_{cov}^{A_f}(S, k) = 1 - \prod_{j=1}^{|S|} (1 - f(s_j, k)) \quad (10)$$

and

$$P_{ove}^{A_f}(S, k) = \frac{\sum_{j=1}^{|S|} f(s_j, k)}{|S|}. \quad (11)$$

Therefore, the coverage and overlap rate of  $A$  can be calculated as

$$C_{cov}^{A_f}(S) = \frac{\sum_{k=1}^{|U|} X_k}{|U|} \quad (12)$$

and

$$C_{ove}^{A_f}(S) = \frac{\sum_{k=1}^{|U|} Y_k}{|U|}, \quad (13)$$

respectively, where

$$X_k = \begin{cases} 1, & P_{cov}^{A_f}(S, k) = 1 \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

and

$$Y_k = \begin{cases} 1, & P_{ove}^{A_f}(S, k) > \frac{1}{|S|} \\ 0, & \text{otherwise} \end{cases} \quad (15)$$

#### 2) Connectivity

In [71], the following theorem of graph theory is used to analyse the network connectivity, which is adopted in this paper for checking the connectivity of the graph induced by  $S$ .

**Theorem 1.** *Let  $G$  be the adjacency matrix of graph induced by  $S$ , and the matrix is expressed as  $M = [m_{ij}]_{i,j=1,2,\dots,|S|} = G + G^2 + \dots + G^{|S|-1}$ . The graph is connected if and only if*

$$m_{ij} \neq 0, \quad \forall i, j \in [1, |S|]. \quad (16)$$

### III. PROBLEM STATEMENT

With the aforementioned models and definitions, in this section we mathematically formulate our problem and prove that this problem is NP-hard.

#### A. Problem Formulation

Given a set of candidate locations  $L$  where SILs can be deployed, as shown in Fig.4, we are interested in selecting the minimum number of candidate locations from  $L$  to deploy SIL nodes while maximizing the overlap rate and meeting the requirements on full coverage and connectivity. This problem is denoted by  $P1$  with the objective functions formulated as

$$\max C_{ove}^{A_f}(S) \quad (17)$$

$$\min |S| \quad (18)$$

s.t.

$$|L| = \rho \sum_{j=1}^{|A|} A_A^j, \quad \forall i \in [1, |S|], s_i \in L \quad (19)$$

$$C_{cov}^{A_f}(S) = 1 \quad (20)$$

$$g_{uv} = \begin{cases} 1, & d(s_u, s_v) \leq \min(R_c^u, R_c^v) \\ 0, & \text{otherwise} \end{cases} \quad (21)$$

$$M = G + G^2 + \dots + G^{|S|-1}, \quad m_{uv} \neq 0$$

$$\forall u, v \in [1, |S|], g_{uv} \in G, m_{uv} \in M, s \in S$$

where (19) means only candidate locations can be used to deploy SIL nodes, and (20) means the solution  $S$  should meet the full coverage requirement. Additionally, (21) means the graph induced by  $S$  is connected, i.e., the network is connected.

#### B. Hardness Analysis

To analyze the hardness of  $P1$ , we need to transform  $P1$  into the known NP-hard problem, e.g., the CSC problem. The following Theorem shows how to prove that  $P1$  is NP-hard.

**Theorem 2.** *The problem  $P1$  is NP-hard.*

*Proof.* The CSC problem is an extension of the Set Cover Problem (SCP), and this problem on complex network graphs, e.g., star and spider, has been proved NP-hard [72], [73]. Given a universal set  $U = \{1, 2, \dots, n\}$ , a family  $F = \{F_1, F_2, \dots, F_m\}$  with  $m$  subsets of  $U$ , a graph  $G$  with vertex set  $V(G) = F$ ,



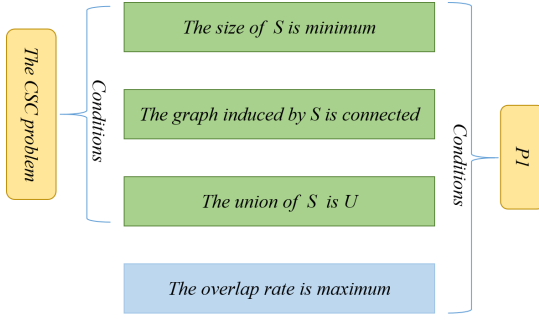


Fig. 7. Illustration of the problem transformation. According to Theorem 2,  $P1$  can be regarded as a CSC problem with an additional constraint on the maximum overlap rate, i.e.,  $P1$  is a special case of the CSC problem.

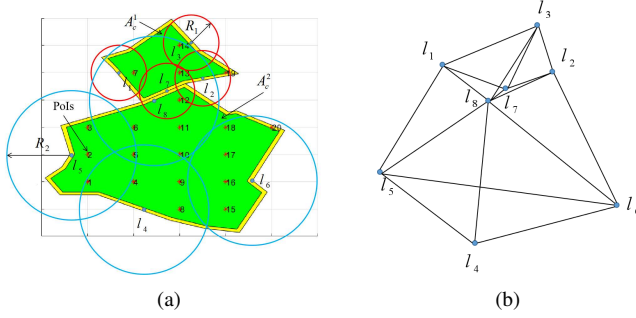


Fig. 8. An instance of  $P1$ : (a) the distribution of candidate locations, with  $U = \{1, 2, \dots, 20\}$ ,  $L = \{l_1, l_2, \dots, l_8\}$ ,  $l_1 = \{7\}$ ,  $l_2 = \{13, 19\}$ ,  $l_3 = \{14\}$ ,  $l_4 = \{1, 4, 5, 8, 9\}$ ,  $l_5 = \{1, 2, 3, 5\}$ ,  $l_6 = \{15, 16, 17, 18, 20\}$ ,  $l_7 = \{12, 13\}$  and  $l_8 = \{5, 6, 7, 10, 11, 12, 13, 14\}$ ; (b) the connected graph constructed by  $L$ . Therefore,  $P1$  is to identify a subset  $S$  out of  $L$ , with the maximum overlap rate, the minimum size of  $S$ , the connected graph induced by  $S$ , and  $U$  being the union of  $S$ .

and the edge set  $E(G)$  consisting edges between some pairs of subsets in  $F$ , the problem is to find a minimum set cover  $C \subseteq F$  to guarantee that the subgraph induced by  $C$  is connected and the union of  $C$  is  $U$ .

For  $P1$ , a set of PoIs in  $A$  corresponds to  $U$  in the CSC problem, and a collection of candidate locations  $L = \{l_1, l_2, \dots, l_{|L|}\}$  corresponds to  $F$ . Each candidate location  $l_i$  is a subset of  $U$  and the graph  $G$  is constructed by  $L$ , where  $V(G) = L$ . Thus,  $P1$  is to identify a subset  $S$  of  $L$  such that all PoIs in  $A$  are covered while the overlap rate is maximised, the size of  $S$ ,  $|S|$ , is minimised, and the graph induced by  $S$  is connected.

Therefore,  $P1$  can be regarded as the CSC problem with an additional constraint on the maximum overlap rate. That is,  $P1$  is a special case of the CSC problem, as shown in Fig. 7. Since the CSC problem on complex network graphs is NP-hard,  $P1$  is NP-hard as well. An instance is given in Fig. 8, where  $U = \{1, 2, \dots, 20\}$  and  $L = \{l_1, l_2, \dots, l_8\}$ .

#### IV. THE HOLE AWARE NODE DEPLOYMENT METHOD

In this section, a SIL node deployment method based on greedy algorithm, referred to as the HANDM, is proposed to solve the defined constrained optimization problem. The HANDM is a two-phase method, in which we first find a set

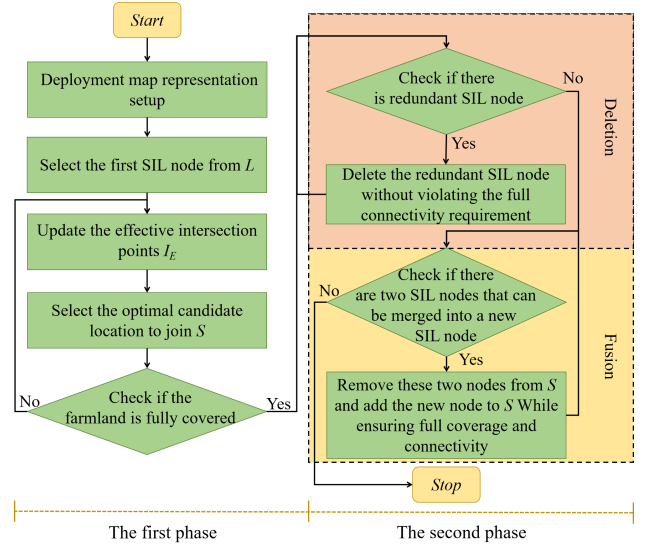


Fig. 9. Flowchart of the proposed HANDM.

of SIL nodes with suboptimal locations. Then, with the aid of deletion and fusion operations, we attempt to finetune these nodes toward the optimal locations for meeting the requirements on full coverage and connectivity, while achieving the maximum overlap rate at the minimum deployment cost. For clarity, the HANDM flowchart is shown in Fig. 9.

##### A. Deployment Map Representation Setup

In this part, the farmland  $A$  is discretized into grid points based on  $\zeta$  so that a set of PoIs  $U$  and a set of BPs  $B$  are determined. The BPs in  $B$  are sorted in clockwise or counterclockwise order. Then, the area of each separate partition is calculated and a set of candidate locations  $L$  is generated, where each PoI is covered by at least one candidate location and there is at least one route between two separate partitions with the common boundary. Finally, according to (8), the maximum transmission distance of each candidate location is calculated.

Note that, as the shape of an actual farmland can be approximated by a polygon, the methods used to identify convexity-concavity of a simple polygon [74], e.g., angling method, left-right-point method, vector-area method, vector-product method, raying method, slopping method and extremity-vertices-order method, can also be used to sort  $B$ . In this paper, we utilize the vector-product method (*Right-Hand Rule*) to sort  $B$  in counterclockwise order. If the reader wants to know more about the process of calculation, we refer them to [74]–[77] for more details.

##### B. Selection of the First SIL Node

According to the core idea of greedy strategy, there are many equivalent initial solutions to the first node in solution space. The works in [41], [47] adopt a simple mechanism for the selection of the first node, i.e., to select the location closest to the regional centre for the first node deployment. However, this mechanism is not reasonable since the deployment strategy caused by this selection mechanism, i.e., the deployment

---

**Algorithm 1** Selection of the First SIL Node.

```

1:  $S \leftarrow \emptyset$ ;  $Max\_Cov \leftarrow 0$ ;
2: Identify MBR based on the polygon's vertices and determine  $CP$  based on MBR;
3: for  $i \in CP$  do
4:   for  $l \in Y(i)$  do
5:     if  $|J(W(l))| == 2$  then
6:       if  $C_{Cov}^{Af}(l) \geq Max\_Cov$  then
7:          $First\_SIL \leftarrow l$ ;
8:          $Max\_Cov \leftarrow C_{Cov}^{Af}(l)$ ;
9:       end if
10:    end if
11:  end for
12: end for
13:  $S = S \cup First\_SIL$ ;
14: Return  $S$ ;

```

To address this issue, we propose a novel strategy to deploy the nodes from RP to the convex point that is on the MBR's boundary paralleled to RB, as shown in Fig. 10(b). A candidate location  $l$  can be selected to deploy the first SIL node only if the following condition is satisfied:

$$l = \arg \max_{l^* \in (L_{CP} \cap L_{JP})} C_{Cov}^{A_f}(l^*) \quad (22)$$

$$L_{CP} = \bigcup_{i \in CP} Y(i) \quad (23)$$

```

1:  $I_B \leftarrow \emptyset$ ;  $I_A \leftarrow \emptyset$ ;  $I_E \leftarrow \emptyset$ ;  $D \leftarrow \emptyset$ ;
2: Calculate the BCV  $V$  based on  $S$ , i.e.,  $V = W(S)$ ;
3: for  $j \in J(V)$  do
4:    $I_B = I_B \cup B(j)$ ;
5: end for
6: for  $m, n \in S, m \neq n$  do
7:   if  $d(m, n) \leq R^m + R^n$  then
8:     Calculate the AIs  $a$ ;
9:      $I_A = I_A \cup a$ ;
10:  end if
11: end for
12:  $I = I_A \cup I_B$ ;
13: for  $i \in I$  do
14:   if  $i$  is within the polygon then
15:     if  $\forall s \in S, d(i, s) \geq R^s$  then
16:        $I_E = I_E \cup i$ ;
17:     end if
18:   end if
19: end for
20: for  $k \in I_E$  do
21:   Calculate the distance  $d_k$  from point  $k$  to RB;
22:    $D = D \cup d_k$ ;
23: end for
24: Sort the  $I_E$  in ascending order in terms of distance;
25: Return  $I_E$ ;

```

$$\forall l'' \in L_{IP}, |J(W(l''))| = 2 \quad (24)$$

The pseudocode of the location selection to deploy the first SIL node in our method is presented in Algorithm 1.

### C. Update of Effective Intersection Points

After adding a new candidate location to  $S$ , the set of effective intersection points  $I_E$  needs to be updated so that each point in  $I_E$  should meet the following condition:

$$d(i, s) \geq R^s, \quad \forall i \in I_E, \quad \forall s \in S, \quad (25)$$

Additionally, to ensure only one coverage hole is generated by  $S$ ,  $I_E$  is sorted in ascending order with respect to the distance from the point in  $I_E$  to the RB. The pseudocode of updating  $I_E$  is presented in Algorithm 2.

For example,  $S = \{s_1, s_2\}$  and  $I_E = \{1, 2, 3\}$  in Fig. 11(a). When  $s_3$  is selected to join  $S$ , as shown in Fig. 11(b),  $I = \{1, 2, \dots, 10\}$  with  $|I| = 10$ . Since the AIs  $\{7, 8\}$  are not within the polygon (the colored area) and the points  $\{5, 6, 9, 10\}$  do not meet (25), there are only four effective intersection points,  $\{1, 2, 3, 4\}$ . It can be observed from Fig. 11(b) that point 1 has the shortest distance to the RB, followed by point 2 and then point 3, finally is point 4. Thus, we have  $I_E = \{1, 2, 3, 4\}$ .

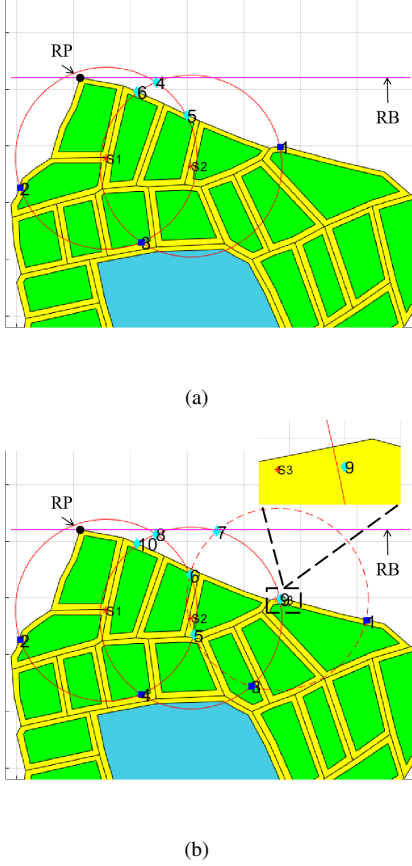


Fig. 11. An instance of updating  $I_E$ : (a)  $S = \{s_1, s_2\}$  and  $I_E = \{1, 2, 3\}$ . (b) Given that  $s_3$  is selected to join  $S$ ,  $I = \{1, 2, \dots, 10\}$  with  $|I| = 10$ . Since the AIs  $\{7, 8\}$  are not within the polygon (the colored area) and the points  $\{5, 6, 9, 10\}$  do not meet (25), there are only four effective intersection points,  $\{1, 2, 3, 4\}$ . Among these four points, the distance from point 1 to RB is the shortest, followed by point 2 and then point 3, finally is point 4. Thus,  $I_E = \{1, 2, 3, 4\}$ .

#### D. Selection of Remaining SIL Nodes

In addition to the first SIL node, the candidate location  $l$  can be selected to join  $S$  in each iteration if it satisfies the following condition:

$$l = \arg \max_{l^* \in (L' \cap L_{JP} \cap L_{Com})} C_{Cov}^{Af}(l^*) \quad (26)$$

s.t.

$$L' = Y(i^*), \quad i^* = \arg \min_{i \in I_E} D(i) \quad (27)$$

$$\forall l'' \in L_{JP}, \quad |J(W(S \cup l''))| = 2 \quad (28)$$

$$L_{Com} = \bigcup_{s \in S} Z(s) \setminus S \quad (29)$$

where  $D(i)$  is a function that returns the distance between the effective intersection point  $i$  and the RB. (26) means the location  $l$  has the maximum coverage rate among the candidates in  $L' \cap L_{JP} \cap L_{Com}$ . (27) means the effective intersection point closest to the RB has the priority to be covered by  $l$ , and (28) means the network built by  $S \cup l$  has only one coverage hole. Additionally, (29) means that the network built by  $S \cup l$  is connected.

#### Algorithm 3 Selection of Remaining SIL Nodes

```

1: while  $C_{Cov}^{Af}(S) \neq 1$  do
2:    $SIL\_Num \leftarrow |S|$ ;
3:    $Max\_Cov \leftarrow C_{Cov}^{Af}(S)$ ;
4:   Apply Algorithm 2 to obtain  $I_E$ ;
5:   while  $SIL\_Num == |S|$  do
6:      $L' = Y(i^*), \quad i^* = \arg \min_{i \in I_E} D(i)$ ;
7:      $T = L' \cap L_{JP} \cap L_{Com}$ ;
8:     if  $|T| \neq 0$  then
9:       for  $t \in T$  do
10:        if  $C_{Cov}^{Af}(S \cup t) \geq Max\_Cov$  then
11:           $Max\_Cov \leftarrow C_{Cov}^{Af}(S \cup t)$ ;
12:           $SIL\_Coor \leftarrow t$ ;
13:        end if
14:      end for
15:       $S = S \cup SIL\_Coor$ ;
16:    else
17:       $I_E \leftarrow I_E \setminus p$ ;
18:    end if
19:  end while
20: end while
21: Return  $S$ ;

```

#### Algorithm 4 Deletion Operation

```

1: for  $s \in S$  do
2:    $S^* \leftarrow S \setminus s$ 
3:   if  $C_{Cov}^{Af}(S^*) == 1$  then
4:     if  $\forall u, v \in [1, |S^*|], m_{uv} \neq 0$  then
5:        $S \leftarrow S^*$ ;
6:     end if
7:   end if
8: end for
9: Return  $S$ ;

```

The pseudocode of selecting the remaining SIL nodes is described in Algorithm 3.

#### E. Optimization of Solution

In the second phase, we design two operations, i.e., deletion and fusion, for optimising  $S$  to meet (17)-(21).

1) *Deletion*: Since  $S$  may contain redundant SIL nodes, these redundant SIL nodes should be deleted without violating the full connectivity requirement. The pseudocode of the deletion operation is described in Algorithm 4.

A redundant SIL node  $s$  can be deleted from  $S$  only if the following condition is satisfied:

$$M = G(S^*) + G(S^*)^2 + \dots + G(S^*)^{|S^*|-1} \quad (30)$$

$$m_{uv} \neq 0$$

$$\forall u, v \in [1, |S^*|], m_{uv} \in M$$

where  $S^* = S \setminus s$ .

For example, in Fig. 12(a), the 10<sup>th</sup>, 11<sup>th</sup>, and 18<sup>th</sup> SIL nodes are three redundant nodes. Since the 11<sup>th</sup> node has no contribution in terms of both coverage and connectivity, it needs to be deleted from  $S$ . However, the 10<sup>th</sup> and 18<sup>th</sup>

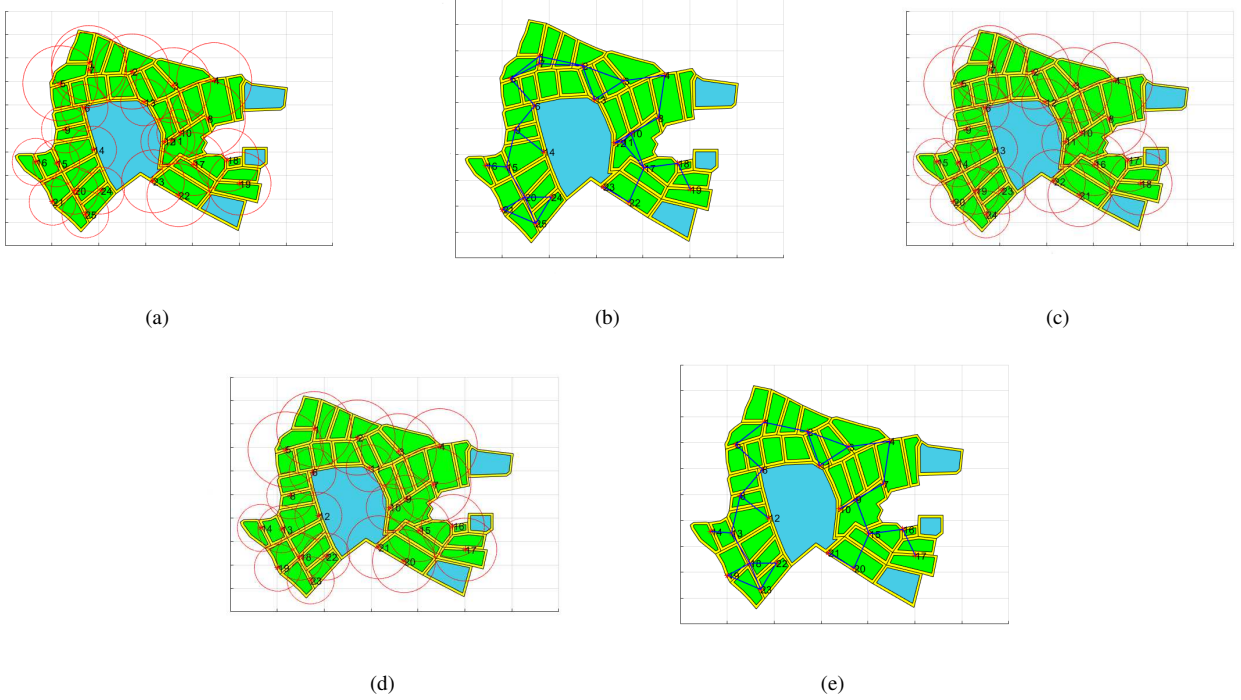


Fig. 12. An illustration of optimization. (a) The deployment result obtained in the first phase. (b) The original network topology. (c) The deployment result with deletion operation. (d) The deployment result with fusion operation. (e) The network topology with fusion operation.

SIL nodes cannot be deleted since the network is disconnected without them, as shown in Fig. 12(b). The deployment result with the deletion operation is shown in Fig. 12(c).

2) *Fusion*: The fusion operation attempts to optimize  $S$  by merging two SIL nodes into a new SIL node, so as to reduce the deployment cost and maximize the overlap rate without violating the requirements on full coverage and connectivity. If the merge is successful, a success flag is returned and the merged SIL node is reflected in  $S$ . Physically, two SIL nodes are removed from  $S$ , and a new SIL node is added to  $S$ . Otherwise, a failure flag is returned and  $S$  remains the same. The pseudocode of fusion operation is described in Algorithm 5.

Let  $(i, j)$  be a pair of SIL nodes with  $d(i, j) \leq R^i + R^j$ . The fusion operation attempts to find a candidate location  $l$  so that the following condition is satisfied:

$$l = \arg \max_{l^* \in (L'' \cap L''')} C_{ove}^{A_f}(S' \cup l^*) \quad (31)$$

s.t.

$$\forall l' \in L'', C_{cov}^{A_f}(S' \cup l') = 1 \quad (32)$$

$$\begin{aligned} & \forall l'' \in L''', \forall u, v \in [1, |S''|], m_{uv} \in M, \\ & M = G(S'') + G(S'')^2 + \dots + G(S'')^{|S''|-1}, \quad (33) \\ & m_{uv} \neq 0 \end{aligned}$$

where  $S' = S \setminus \{i, j\}$ ,  $S'' = S' \cup l''$ , and  $(L'', L''') \subset L \setminus S$ .

For example, the coverage area of the 6<sup>th</sup> and 7<sup>th</sup> SIL nodes in Fig. 12(c) can be covered by a new SIL node, i.e., the 6<sup>th</sup> node in Fig. 12(d). Thus, the 6<sup>th</sup> and 7<sup>th</sup> nodes in Fig. 12(c) are removed from  $S$ , and the 6<sup>th</sup> node in Fig. 12(d) is added to

#### Algorithm 5 Fusion Operation

---

```

1: for  $i, j \in S, i \neq j, d(i, j) \leq R^i + R^j$  do
2:    $S' \leftarrow S \setminus \{i, j\}; Max\_Ove \leftarrow 0; Flag \leftarrow 0;$ 
3:   for  $l \in L \setminus S$  do
4:      $S^* \leftarrow S' \cup l;$ 
5:     if  $C_{cov}^{A_f}(S^*) == 1$  then
6:       if  $\forall u, v \in [1, |S^*|], m_{uv} \neq 0$  then
7:         if  $C_{ove}^{A_f}(S^*) \geq Max\_Ove$  then
8:            $S'' = S^*;$ 
9:            $Max\_Ove = C_{ove}^{A_f}(S^*);$ 
10:           $Flag \leftarrow 1;$ 
11:        end if
12:      end if
13:    end if
14:  end for
15:  if  $Flag == 1$  then
16:     $S \leftarrow S'';$ 
17:  end if
18: end for
19: Return  $S;$ 

```

---

$S$ . Consequently, the network topology with fusion operation is shown in Fig. 12(e).

## V. PERFORMANCE EVALUATION

In this section, the efficiency and effectiveness of the proposed method for solving the cSILDLP are evaluated in terms of deployment cost and Overlap Rate ( $OR$ ). We compare our proposed HANDM with four benchmark algorithms: the growth rings like deployment algorithm (GRLD) [34], the

TABLE IV  
DEPLOYMENT STRATEGY COMPARISONS BETWEEN BENCHMARK  
ALGORITHMS AND OUR PROPOSED HANDM

Method	Deployment Strategy
HANDM	SIL nodes are deployed from the RP to the convex point that is on the MBR's boundary paralleled to the RB.
GRLD	SIL nodes are deployed round by round from the boundary to the centre until the entire farmland is completely covered.
GAA	SIL nodes are deployed from the centre to the periphery. Meanwhile, the new SIL node added in each iteration must satisfy two conditions: 1) at least one effective intersection point is covered; 2) the maximum coverage is achieved.
CLBGA	SIL nodes are deployed from the centre to the periphery. The SIL node with the maximum coverage is selected in each iteration.
R-Gr	CLBGA is repeated for $p_{size}$ times with different initial solutions, and the best one among $p_{size}$ solutions is selected.

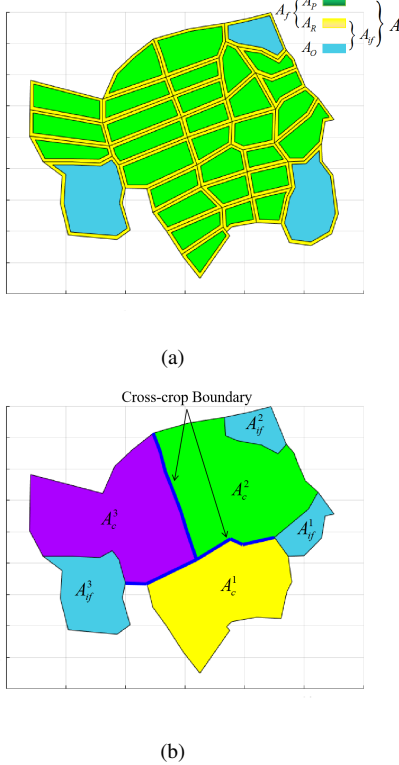


Fig. 13. Map 2: (a) the diagram, (b) the distribution of mono-crop plots.

greedy approximate algorithm (GGA) [41], the candidate location based greedy algorithm (CLBGA) [43], and the randomized greedy algorithm (R-Gr) [46]. The GRLD represents a group of methods to deploy nodes from the boundary to the centre, while the other three methods deploy nodes from the centre to the periphery. The main differences between the proposed HANDM and these four peer algorithms in terms of deployment strategy are listed in Table IV-E.

#### A. Simulation Setup

Since the original GRLD is a deterministic deployment algorithm, which cannot be directly applied to solve the

TABLE V  
SIMULATION PARAMETERS

Parameter	Value	Parameter	Value
$P$	0(dbm)	$m$ (Maps 1, 2)	700/600(m)
$\zeta$	1(m)	$n$ (Maps 1, 2)	500/450(m)
$p_{size}$	30	$A_A$ (Maps 1, 2)	13.8/13.6(hm <sup>2</sup> )
$\rho$	$4 \times 10^{-3}$	$\gamma_{th}$	$[2, 10] \times 10^{-8}$ (Steps of $2 \times 10^{-8}$ )

TABLE VI  
THREE DESIGN SCENARIOS.

	Scenario 1			Scenario 2			Scenario 3		
	$A_c^1$	$A_c^2$	$A_c^3$	$A_c^1$	$A_c^2$	$A_c^3$	$A_c^1$	$A_c^2$	$A_c^3$
$R$	100	80	71	80	50	67	100	80	71
$\alpha$	3.66	3.85	3.71	3.66	3.85	3.71	3.76	3.96	3.83

cSILDP, we adopt the core idea of this algorithm to deploy nodes round by round from the boundary to the centre of the monitoring area until the entire area is fully covered. In other words, we use the minimum number of SIL nodes to fully cover the boundary of each round until the full coverage is achieved. Additionally, to enhance the diversity of population,  $p_{size}$  initial solutions of the R-Gr are randomly selected from  $L$ .

In addition to the map given in Fig. 4, another map shown in Fig. 13 is used in our simulations for further comparisons. The simulation parameters are listed in Table V. Moreover, Table VI lists three design scenarios, in which scenarios 1 and 2 are the same in  $\alpha$  but different in  $R$ , while scenarios 1 and 3 are the same in  $R$  but different in  $\alpha$ . It is worth noting that the values of these two parameters are set based on the previous works that mentioned in Sections *Effective Killing Distance* and *Path Loss Model*. For example, the path loss exponent for rice at maturity stage is about 3.66 [61], and the effective killing distance for *Ostrinia furnacalis* is about 100m [63]. For the reader who wants to know more about specific parameters, please refer to the papers mentioned in these two Sections.

The simulations are performed by MATLAB R2018b on a PC with 64-bit Microsoft Windows 10 operating system, 8GB RAM and 2.2 GHz-Core i7 CPU. All experimental values are achieved over 30 independent simulations. For each simulation, candidate locations where SILs can be deployed are randomly generated according to a uniform distribution.

#### B. Simulation Results

In this part, the ability of HANDM to offer efficient solution to the cSILDP is illustrated. Furthermore, the influence of the receiver sensitivity threshold  $\gamma_{th}$  on the deployment cost and overlap rate is evaluated as well.



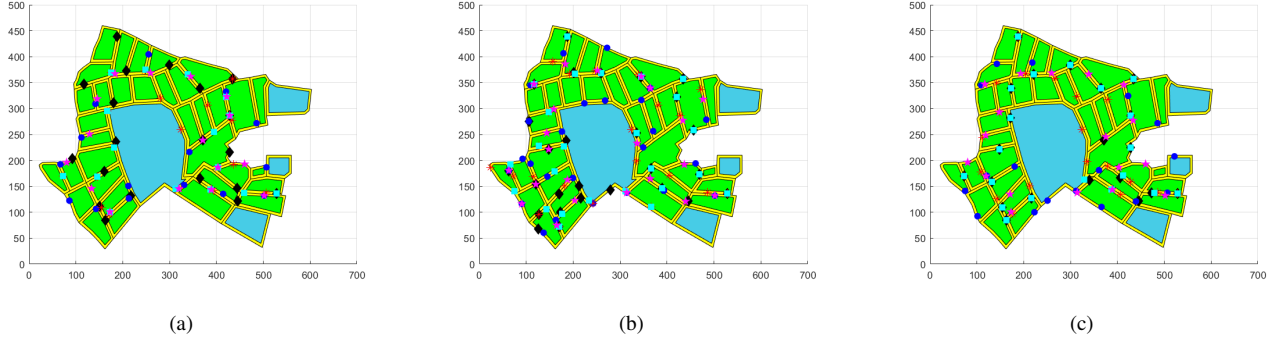


Fig. 14. Implementation of the SIL node deployment with receiver sensitivity threshold  $\gamma_{th} = 6 \times 10^{-8}$  for Map 1 by CLBGA (black diamonds), by R-Gr (cyan squares), by GAA (red stars), by GRLD (blue dots), and by HANDM (pink pentagrams): (a) in Scenario 1, (b) in Scenario 2, (c) in Scenario 3.

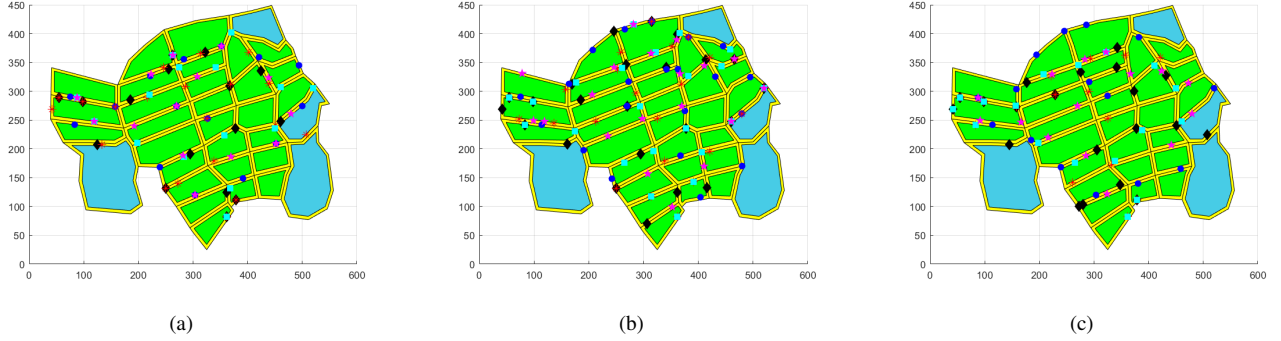


Fig. 15. Implementation of the SIL node deployment with receiver sensitivity threshold  $\gamma_{th} = 6 \times 10^{-8}$  for Map 2 by CLBGA (black diamonds), by R-Gr (cyan squares), by GAA (red stars), by GRLD (blue dots), and by HANDM (pink pentagrams): (a) in Scenario 1, (b) in Scenario 2, (c) in Scenario 3.

TABLE VII  
MAIN PARAMETERS IN FIGS. 14 AND 15

Method	Map 1						Map 2					
	Scenario 1		Scenario 2		Scenario 3		Scenario 1		Scenario 2		Scenario 3	
	$ S $	$OR(\%)$	$ S $	$OR(\%)$	$ S $	$OR(\%)$	$ S $	$OR(\%)$	$ S $	$OR(\%)$	$ S $	$OR(\%)$
HANDM	16	87.97%	22	78.53%	18	92.05%	14	87.14%	19	72.29%	15	87%
GRLD	17	88%	26	69.41%	20	91.14%	15	88.28%	22	70.7%	17	88.83%
R-Gr	16	87.69%	25	82.08%	20	95.82%	15	90.99%	22	84.86%	17	93.46%
CLBGA	20	91.87%	27	84.84%	21	95.77%	16	85.54%	23	85.31%	20	91.97%
GAA	17	87.1%	25	78.3%	22	97.76%	20	95.83%	23	80.17%	17	86.97%

Figs. 14 and 15 plot the implementation of the SIL node deployments with the receiver sensitivity threshold  $\gamma_{th} = 6 \times 10^{-8}$  for the five deployment methods under study in three scenarios. The parameters are listed in Table VII. It can be observed that the deployment cost increases with the decrease in  $R$  and the increase in  $\alpha$ . We also notice that our proposed method performs much better than the other methods in terms of the deployment cost for all the three scenarios. When the deployment cost of our method is the same as that of the other methods, the overlap rate of our method is slightly better than that of the others.

Next, we compare the five deployment methods with different values of  $\gamma_{th}$ . Figs. 16 and 17 plot the average deployment cost, with  $\gamma_{th}$  increasing from  $2 \times 10^{-8}$  to  $10^{-7}$  by a step  $2 \times 10^{-8}$ . The experimental results of 30 independent simu-

lations for Maps 1 and 2 are shown in Figs. 18 and 19. The average  $OR$  of the five deployment methods is listed in Table VIII. To facilitate the comparisons on  $OR$ , the normalization of overlap rate,  $NOR$ , is used to represent the ratio of overlap rate to deployment cost. The results of normalization for Maps 1 and 2 are presented in Figs. 20 and 21. It can be seen that the deployment cost is increased as the receiver sensitivity threshold increases. Moreover, our proposed method performs much better than the other methods in terms of deployment cost. Generally, the  $NOR$  of our method is better than that of the others. In some special instances, the  $NOR$  of our method is lower than that of the other, e.g.,  $NOR$  of the R-Gr is higher than that of the HANDM when  $\gamma_{th} = 8 \times 10^{-8}$  in Fig. 20(c), where the deployment cost of R-Gr is however higher than that of our HANDM.

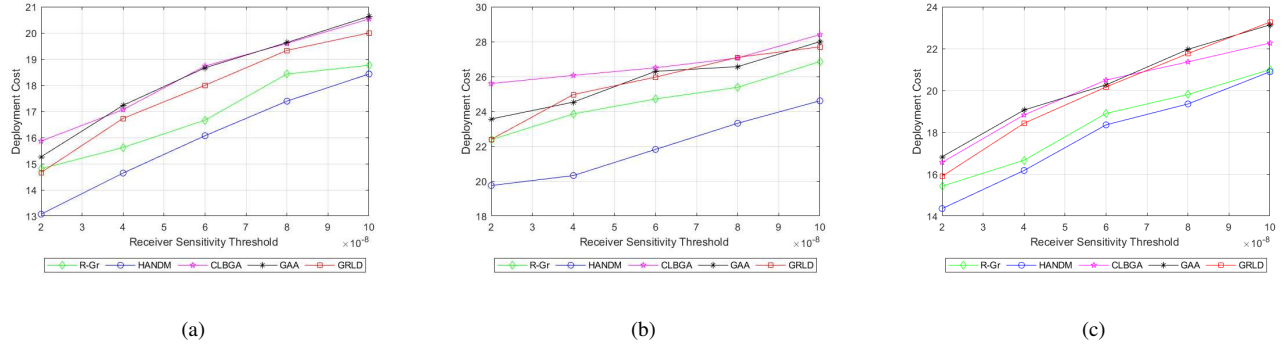


Fig. 16. Average deployment cost comparisons for Map 1: (a) in Scenario 1, (b) in Scenario 2, (c) in Scenario 3.

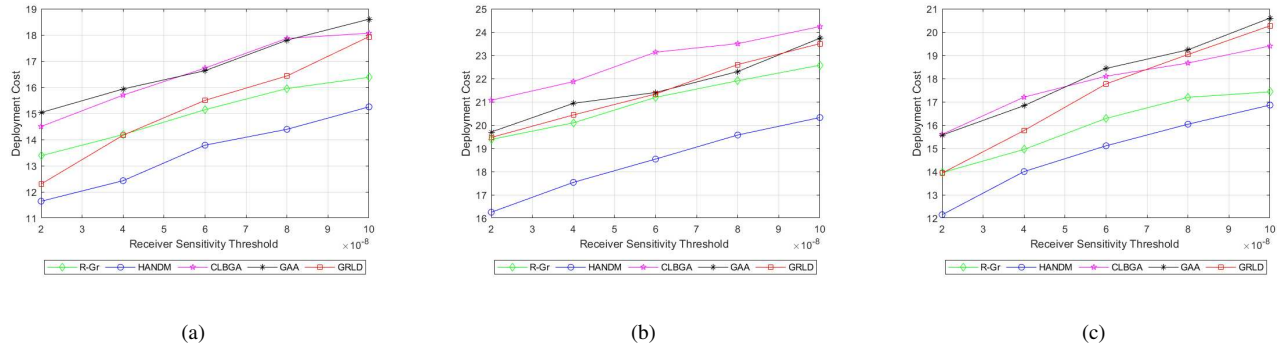


Fig. 17. Average deployment cost comparisons for Map 2: (a) in Scenario 1, (b) in Scenario 2, (c) in Scenario 3.

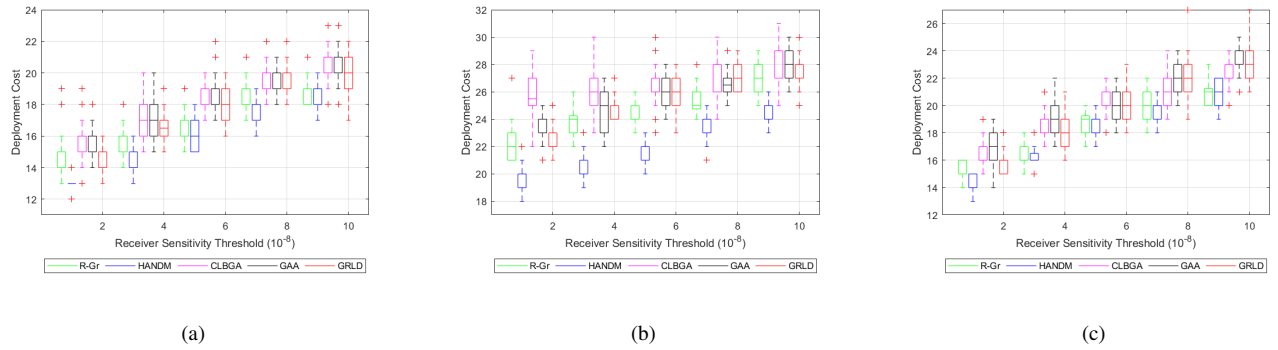


Fig. 18. Boxplot of deployment cost for Map 1: (a) in Scenario 1, (b) in Scenario 2, (c) in Scenario 3.

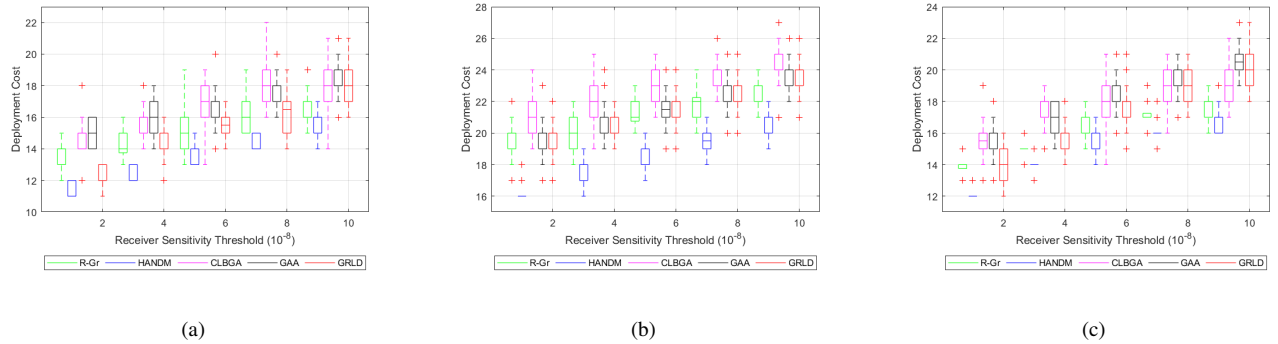
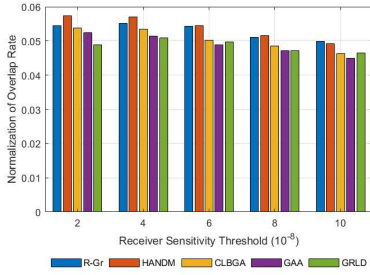


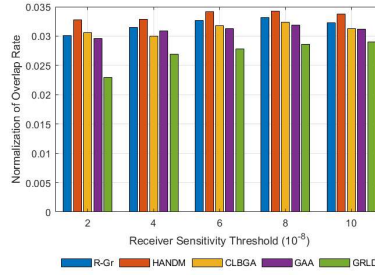
Fig. 19. Boxplot of deployment cost for Map 2: (a) in Scenario 1, (b) in Scenario 2, (c) in Scenario 3.

TABLE VIII  
EXPERIMENTAL RESULTS.

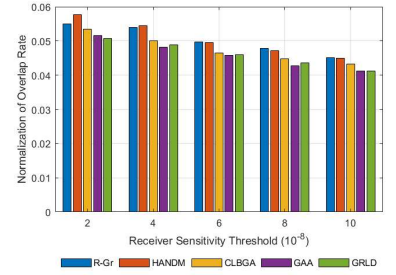
$\gamma_{th}$ ( $10^{-8}$ )	Method	Map 1						Map 2					
		Scenario 1		Scenario 2		Scenario 3		Scenario 1		Scenario 2		Scenario 3	
		S	OR(%)	S	OR(%)	S	OR(%)	S	OR(%)	S	OR(%)	S	OR(%)
2	HANDM	13.07	75.07%	19.75	64.79%	14.35	82.92%	11.64	75.74%	16.25	61.36%	12.14	80.54%
	GRLD	14.66	71.68%	22.4	51.31%	15.9	80.56%	12.3	67.66%	19.46	60.87%	13.93	79.02%
	R-Gr	14.8	80.79%	22.38	67.39%	15.42	84.71%	13.38	76.68%	19.38	69.73%	13.95	81.21%
	CLBGA	15.86	85.42%	25.6	78.43%	16.56	88.68%	14.5	78%	21.06	77.19%	15.6	86.72%
	GAA	15.26	79.93%	23.56	69.71%	16.83	86.95%	15.03	83.23%	19.7	70.23%	15.56	86.1%
4	HANDM	14.64	83.55%	20.32	66.83%	16.17	88.06%	12.42	81%	17.53	70.2%	14	86.87%
	GRLD	16.73	85.06%	24.96	67.11%	18.43	89.92%	14.16	80.09%	20.43	67.56%	15.76	89.23%
	R-Gr	15.61	86.27%	23.85	75.01%	16.66	89.99%	14.19	82.71%	20.09	74.65%	14.95	86.79%
	CLBGA	17.06	91.3%	26.06	78.22%	18.83	94.11%	15.7	86.68%	21.86	78.63%	17.2	91%
	GAA	17.23	88.53%	24.53	75.84%	19.06	91.87%	15.93	87.33%	20.93	75.77%	16.83	88.86%
6	HANDM	16.07	87.64%	21.82	74.62%	18.35	91.03%	13.78	86.37%	18.53	74.89%	15.1	88.49%
	GRLD	18	89.36%	25.96	72.19%	20.16	92.82%	15.5	87.29%	21.33	72.93%	17.76	91.86%
	R-Gr	16.66	90.58%	24.71	80.76%	18.9	94.1%	15.14	88.46%	21.19	80.02%	16.28	90.86%
	CLBGA	18.73	94.19%	26.5	84.13%	20.5	95.14%	16.73	89.82%	23.13	83.9%	18.1	91.95%
	GAA	18.66	91.33%	26.3	82.14%	20.26	92.86%	16.63	88.43%	21.4	78.6%	18.43	92.16%
8	HANDM	17.39	89.86%	23.32	79.79%	19.35	91.38%	14.39	87.29%	19.57	78.93%	16.03	89.91%
	GRLD	19.33	91.26%	27.1	77.48%	21.76	94.89%	16.43	89.86%	22.6	79.8%	19.03	93.33%
	R-Gr	18.42	93.97%	25.38	84.26%	19.8	94.63%	15.95	89.94%	21.9	82.64%	17.19	91.85%
	CLBGA	19.6	94.95%	27.06	87.51%	21.36	95.71%	17.86	91.53%	23.5	86.73%	18.66	93.51%
	GAA	19.63	92.54%	26.56	84.8%	21.96	93.99%	17.8	91.48%	22.3	82.55%	19.23	93.44%
10	HANDM	18.42	90.63%	24.6	83.13%	20.89	93.93%	15.25	88.68%	20.32	82.23%	16.85	90.65%
	GRLD	20	93.01%	27.7	80.36%	23.26	95.7%	17.93	92.36%	23.5	82.78%	20.26	94.6%
	R-Gr	18.76	93.46%	26.85	86.73%	21	94.73%	16.38	90.9%	22.57	85.56%	17.42	92.36%
	CLBGA	20.53	95.17%	28.4	88.95%	22.26	96.4%	18.06	92.39%	24.23	88.19%	19.4	94.7%
	GAA	20.63	92.85%	28	87.28%	23.13	95.24%	18.6	92.56%	23.73	86.86%	20.6	94.2%



(a)

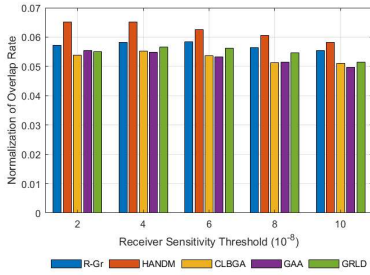


(b)

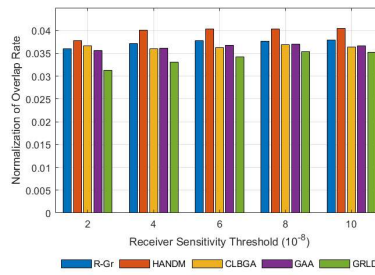


(c)

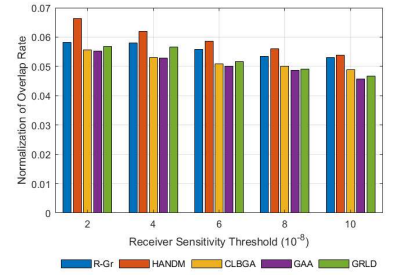
Fig. 20. *NOR* comparisons for Map 1: (a) in Scenario 1, (b) in Scenario 2, (c) in Scenario 3.



(a)



(b)



(c)

Fig. 21. *NOR* comparisons for Map 2: (a) in Scenario 1, (b) in Scenario 2, (c) in Scenario 3.

Based on the above simulation results, the conclusions can be drawn as the following:

- 1) The deployment strategy adopted in GAA and CLBGA cannot effectively deal with the cSILDP in the RoI with an irregular shape, especially for the irregular RoI with more convex vertices. Therefore, our HANDM has a better performance than these two methods in terms of deployment cost. Although the R-Gr can reduce the impact of initial solution on the dimension of solution vector, the performance of HANDM is still better than that of the R-Gr in terms of deployment cost since the latter lacks the optimization in the solution.
- 2) The deployment strategy of GRLD can eliminate the impact of irregular RoI on the dimension of solution vector, but due to the lack of optimization in the solution, it is difficult for GRLD to reduce the dimension of solution vector. This is the reason why the HANDM outperforms the GRLD in terms of deployment cost.

## VI. CONCLUSION AND DISCUSSION

### A. Summary

As there are some physical constraints on the deployment of SIL nodes, found in our previous work on the SILDP, we have further investigated the cSILDP in a mixed-crop farmland where the deployment of SIL nodes is restricted to a set of candidate locations that are on the ridges. We have formulated the cSILDP as a CSC problem and proven that it is NP-hard. Since the number of coverage holes is highly related with the dimension of solution vector, the number of coverage holes generated in each iteration need to be reduced as much as possible. Motivated by this, we have proposed a two-phase deployment algorithm, referred to as HANDM, to solve the presented problem. In the first phase, we have proposed a novel deployment strategy to ensure that there is only one coverage hole in each iteration. On the basis of this strategy, a set of SIL nodes is deployed at suboptimal locations. In the second phase, with the aid of deletion and fusion operations, we have finetuned the deployed nodes toward the optimal locations without violating the full coverage and connectivity requirements. Illustrative simulations have been carried out to evaluate the proposed method and verify its effectiveness in terms of deployment cost and overlap rate.

However, since this paper focuses on the minimisation of deployment cost without violating full coverage and connectivity requirements, the network formed by the proposed method is simple, i.e., 1-connectivity. Those problems that affect the network connectivity, e.g., asymmetric link caused by irregular transmission range of a SIL node and the node fault caused by solar panel or wireless module, will lead to the anomalies, e.g., communication anomaly and system anomaly [78], [79]. Therefore, the proposed method cannot guarantee a better tolerance to network faults than other types of IoTs.

### B. Future Work

We plan to carry out our future work in the following three aspects:

- 1) As we focus on the cSILDP with 1-connectivity in this paper, the network is disconnected if any communication link is interrupted. Since the transmit power  $P$  is an important parameter that affects the transmission distance of a SIL node, we will further study the dynamic allocation of transmit power over SIL nodes to enhance the network fault tolerance.
- 2) Due to the dynamic agricultural environment, SIL nodes have the same faults as other IoT devices. If there is a damaged SIL node, the network formed by the proposed method may be disconnected. Therefore, how to design a deployment mechanism that can provide better fault tolerance at the minimum deployment cost without violating the requirements of full coverage and connectivity is another work in the future.
- 3) The Long Range (LoRa) protocol is one of the most successful technologies in Low Power Wide Area Networks (LPWANS), since it enables robust long-distance and low-power communications which are suitable for IoT applications. Therefore, we will solve the cSILDP with LoRa to further improve the network reliability.
- 4) Since the use of mobile agents can support a more resilient deployment [80]–[82], how to use UAVs to optimize the network robustness against the anomalies is another problem we plan to solve.

## REFERENCES

- [1] K. Huang, K. Li, L. Shu, X. Yang, T. Gordon, and X. Wang, "High voltage discharge exhibits severe effect on zigbee-based device in solar insecticidal lamps internet of things," *IEEE Wireless Communications*, pp. 1–6, 2020.
- [2] K. Huang, K. Li, L. Shu, and X. Yang, "Demo abstract: High voltage discharge exhibits severe effect on zigbee-based device in solar insecticidal lamps internet of things," in *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 1266–1267, 2020.
- [3] X. Yang, L. Shu, J. Chen, M. A. Ferrag, J. Wu, E. Nurellari, and K. Huang, "A survey on smart agriculture: Development modes, technologies, and security and privacy challenges," *IEEE/CAA Journal of Automatica Sinica*, pp. 1–30, 2020.
- [4] X. Yang, L. Shu, K. Huang, K. Li, Z. Huo, Y. Wang, X. Wang, Q. Lu, and Y. Zhang, "Characteristics analysis and challenges for fault diagnosis in solar insecticidal lamps internet of things," *Smart Agriculture*, no. 2, pp. 11–27, 2020. (In Chinese with English abstract).
- [5] G. Hu, "Pest automatic counting and publishing system based on pest-killing lamp's voltage disturbance signals," Master's thesis, Nanjing Agricultural University, 2017.
- [6] K. Li, L. Shu, K. Huang, Y. Sun, F. Yang, Y. Zhang, Z. Huo, Y. Wang, X. Wang, Q. Lu, and Y. Zhang, "Research and prospect of solar insecticidal lamps internet of things," *Smart Agriculture*, vol. 1, no. 3, pp. 13–28, 2019. (In Chinese with English abstract).
- [7] F. Subhan, M. Noreen, M. Imran, M. Tariq, A. Khan, and M. Shoaib, "Impact of node deployment and routing for protection of critical infrastructures," *IEEE Access*, vol. 7, pp. 11502–11514, 2019.
- [8] X. Fu, G. Fortino, P. Pace, G. Aloï, and W. Li, "Environment-fusion multipath routing protocol for wireless sensor networks," *Information Fusion*, vol. 53, pp. 4–19, 2020.
- [9] X. Fu, G. Fortino, W. Li, P. Pace, and Y. Yang, "Wsns-assisted opportunistic network for low-latency message forwarding in sparse settings," *Future Generation Computer Systems*, vol. 91, pp. 223–237, 2019.
- [10] L. Shu, J. Xiong, L. Wang, J. Niu, and Q. Wang, "Poster abstract: Studied wind sensor nodes deployment towards accurate data fusion for ship movement controlling," in *2013 ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pp. 339–340, 2013.
- [11] X. Fu, P. Pace, G. Aloï, L. Yang, and G. Fortino, "Topology optimization against cascading failures on wireless sensor networks using a memetic algorithm," *Computer Networks*, vol. 177, p. 107327, 2020.

- [12] D. L. Ndzi, A. Harun, F. M. Ramli, M. L. Kamarudin, A. Zakaria, A. Y. M. Shakaff, M. N. Jaafar, S. Zhou, and R. S. Farook, "Wireless sensor network coverage measurement and planning in mixed crop farming," *Computers & Electronics in Agriculture*, vol. 105, pp. 83–94, 2014.
- [13] R. Elhabyan, W. Shi, and M. St-Hilaire, "Coverage protocols for wireless sensor networks: Review and future directions," *Journal of Communications and Networks*, vol. 21, no. 1, pp. 45–60, 2019.
- [14] P. Si, S. Wang, L. Shu, R. Ning, and Z. Fu, "Optimal deployment for target-barrier coverage problems in wireless sensor networks," *IEEE Systems Journal*, 2020. doi=10.1109/JSYST.2020.2990395.
- [15] P. Si, J. Ma, F. Tao, Z. Fu, and L. Shu, "Energy-efficient barrier coverage with probabilistic sensors in wireless sensor networks," *IEEE Sensors Journal*, vol. 20, no. 10, pp. 5624–5633, 2020.
- [16] S. He, J. Chen, X. Li, X. Shen, and Y. Sun, "Mobility and intruder prior information improving the barrier coverage of sparse sensor networks," *IEEE Transactions on Mobile Computing*, vol. 13, no. 6, pp. 1268–1282, 2014.
- [17] X. Zhu, J. Li, and M. Zhou, "Optimal deployment of energy-harvesting directional sensor networks for target coverage," *IEEE Systems Journal*, vol. 13, no. 1, pp. 377–388, 2019.
- [18] A. A. Obinikpo, Y. Zhang, H. Song, T. H. Luan, and B. Kantarci, "Queueing algorithm for effective target coverage in mobile crowd sensing," *IEEE Internet of Things Journal*, vol. 4, no. 4, pp. 1046–1055, 2017.
- [19] X. Liu, "A deployment strategy for multiple types of requirements in wireless sensor networks," *IEEE Transactions on Cybernetics*, vol. 45, pp. 2364–2376, Oct 2015.
- [20] F. Yang, L. Shu, K. Huang, K. Li, G. Han, and Y. Liu, "A partition-based node deployment strategy in solar insecticidal lamps internet of things," *IEEE Internet of Things Journal*, vol. 7, no. 11, pp. 11223–11237, 2020.
- [21] L. Dai, B. Wang, L. T. Yang, X. Deng, and L. Yi, "A nature-inspired node deployment strategy for connected confident information coverage in industrial internet of things," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 9217–9225, 2019.
- [22] H. Wu and M. Shahidehpour, "Applications of wireless sensor networks for area coverage in microgrids," *IEEE Transactions on Smart Grid*, vol. 9, no. 3, pp. 1590–1598, 2018.
- [23] M. Farsi, M. A. Elhousseini, M. Badawy, H. Arafat Ali, and H. Zain Eldin, "Deployment techniques in wireless sensor networks, coverage and connectivity: A survey," *IEEE Access*, vol. 7, pp. 28940–28954, 2019.
- [24] Z. Fei, B. Li, S. Yang, C. Xing, H. Chen, and L. Hanzo, "A survey of multi-objective optimization in wireless sensor networks: Metrics, algorithms, and open problems," *IEEE Communications Surveys Tutorials*, vol. 19, no. 1, pp. 550–586, 2017.
- [25] C. Tsai, P. Tsai, J. Pan, and H. Chao, "Metaheuristics for the deployment problem of WSN: A review," *Microprocessors & Microsystems*, vol. 39, no. 8, pp. 1305–1317, 2015.
- [26] D. S. Deif and Y. Gadallah, "Classification of wireless sensor networks deployment techniques," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 2, pp. 834–855, 2014.
- [27] W. Bang, "Coverage problems in sensor networks: A survey," *ACM Computing Surveys*, vol. 43, no. 4, pp. 1–56, 2011.
- [28] R. Kershner, "The number of circles covering a set," *Americal Journal of Mathematics*, vol. 61, no. 3, pp. 665–671, 1939.
- [29] B. Wang, H. Xu, W. Liu, and L. T. Yang, "The optimal node placement for long belt coverage in wireless networks," *IEEE Transactions on Computers*, vol. 64, no. 2, pp. 587–592, 2015.
- [30] B. Wang, X. Deng, W. Liu, L. T. Yang, and H. Chao, "Confident information coverage in sensor networks for field reconstruction," *IEEE Wireless Communications*, vol. 20, no. 6, pp. 74–81, 2013.
- [31] C. Chang, C. Chang, Y. Chen, and H. Chang, "Obstacle-resistant deployment algorithms for wireless sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 6, pp. 2925–2941, 2009.
- [32] B. Wang, H. Xu, W. Liu, and H. Liang, "A novel node placement for long belt coverage in wireless networks," *IEEE Transactions on Computers*, vol. 62, no. 12, pp. 2341–2353, 2013.
- [33] C. Bhattacharyya and S. Bhattacharya, "Ldm (layered deployment model): A novel framework to deploy sensors in an irregular terrain," *Wireless Sensor Network*, vol. 3, pp. 189–197, Jan 2011.
- [34] Z. Liao, J. Wang, S. Zhang, and X. Zhang, "A deterministic sensor placement scheme for full coverage and connectivity without boundary effect in wireless sensor networks," *Ad-Hoc and Sensor Wireless Networks*, vol. 19, pp. 327–351, 2013.
- [35] I. Khoufi, P. Minet, A. Laouiti, and E. Livolant, "A simple method for the deployment of wireless sensors to ensure full coverage of an irregular area with obstacles," in *Proceedings of the 17th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, pp. 203–210, 2014.
- [36] Y. Yoon and Y. Kim, "An efficient genetic algorithm for maximum coverage deployment in wireless sensor networks," *IEEE Transactions on Cybernetics*, vol. 43, no. 5, pp. 1473–1483, 2013.
- [37] Z. Wang, H. Xie, D. He, and S. Chan, "Wireless sensor network deployment optimization based on two flower pollination algorithms," *IEEE Access*, vol. 7, pp. 180590–180608, 2019.
- [38] J. Seok, J. Lee, W. Kim, and J. Lee, "A bipopulation-based evolutionary algorithm for solving full area coverage problems," *IEEE Sensors Journal*, vol. 13, no. 12, pp. 4796–4807, 2013.
- [39] O. M. Alia and A. Al-Ajour, "Maximizing wireless sensor network coverage with minimum cost using harmony search algorithm," *IEEE Sensors Journal*, vol. 17, no. 3, pp. 882–896, 2017.
- [40] C. Yang and K. Chin, "On nodes placement in energy harvesting wireless sensor networks for coverage and connectivity," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 1, pp. 27–36, 2017.
- [41] Q. Tang, X. Wang, G. Tong, and X. Li, "Studying on minimal cover set node deployment in wireless sensor networks," in *2015 World Symposium on Computer Networks and Information Security (WSCNIS)*, pp. 1–7, 2015.
- [42] H. Mostafaei and M. S. Obaidat, "A greedy overlap-based algorithm for partial coverage of heterogeneous wsns," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, pp. 1–6, 2017.
- [43] J. Zhu and B. Wang, "Sensor placement algorithms for confident information coverage in wireless sensor networks," in *2014 23rd International Conference on Computer Communication and Networks (ICCCN)*, pp. 1–7, 2014.
- [44] M. Rebai, M. L. Berre, H. Snoussi, F. Hnaïen, and L. Khoukhi, "Sensor deployment optimization methods to achieve both coverage and connectivity in wireless sensor networks," *Computers & Operations Research*, vol. 59, pp. 11–21, 2015.
- [45] L. Dai and B. Wang, "Sensor placement based on delaunay triangulation for complete confident information coverage in an area with obstacles," in *2015 IEEE 34th International Performance Computing and Communications Conference (IPCCC)*, pp. 1–8, 2015.
- [46] T. Grossman and A. Wool, "Computational experience with approximation algorithms for the set covering problem," *European Journal of Operational Research*, vol. 101, no. 1, pp. 81–92, 1997.
- [47] K. Lu and H. Sun, "Greedy approximation algorithm of minimum cover set in wireless sensor networks," *Journal of Software*, vol. 21, no. 10, pp. 2656–2665, 2010. (In Chinese with English abstract).
- [48] J. Jiang, L. Fang, H. Zhang, and W. Dou, "An algorithm for minimal connected cover set problem in wireless sensor networks," *Journal of Software*, vol. 17, no. 17, pp. 175–184, 2006. (In Chinese with English abstract).
- [49] N. Liu, W. Cao, Y. Zhu, J. Zhang, F. Pang, and J. Ni, "Node deployment with  $k$ -connectivity in sensor networks for crop information full coverage monitoring," *Sensors*, vol. 16, no. 12, 2016.
- [50] H. M. Jawad, A. M. Jawad, R. Nordin, S. K. Gharghan, N. F. Abdullah, M. Ismail, and M. J. Abu-AlShaer, "Accurate empirical path-loss model based on particle swarm optimization for wireless sensor networks in smart agriculture," *IEEE Sensors Journal*, vol. 20, no. 1, pp. 552–561, 2020.
- [51] H. Yao, K. Lin, I. Ungurean, and Y. Yang, "Constrained relay node deployment in wireless sensor network for structural health monitoring," in *2019 International Conference on Sensing and Instrumentation in IoT Era (ISSI)*, pp. 1–5, 2019.
- [52] S. Misra, S. D. Hong, G. Xue, and J. Tang, "Constrained relay node placement in wireless sensor networks: Formulation and approximations," *IEEE/ACM Transactions on Networking*, vol. 18, no. 2, pp. 434–447, 2010.
- [53] A. Chelli, M. Bagaa, D. Djenouri, I. Balasingham, and T. Taleb, "One-step approach for two-tiered constrained relay node placement in wireless sensor networks," *IEEE Wireless Communications Letters*, vol. 5, no. 4, pp. 448–451, 2016.
- [54] S. Misra, N. E. Majd, and H. Huang, "Approximation algorithms for constrained relay node placement in energy harvesting wireless sensor networks," *IEEE Transactions on Computers*, vol. 63, no. 12, pp. 2933–2947, 2014.
- [55] D. Yang, S. Misra, X. Fang, G. Xue, and J. Zhang, "Two-tiered constrained relay node placement in wireless sensor networks: Computational complexity and efficient approximations," *IEEE Transactions on Mobile Computing*, vol. 11, no. 8, pp. 1399–1411, 2012.



- [56] C. Ma, W. Liang, and M. Zheng, "Lifetime constrained relay node placement in wsns: A cluster-based approximation algorithm," in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, pp. 1–5, 2017.
- [57] D. Djenouri and M. Bagaa, "Energy-aware constrained relay node deployment for sustainable wireless sensor networks," *IEEE Transactions on Sustainable Computing*, vol. 2, no. 1, pp. 30–42, 2017.
- [58] S. Li, H. Gao, and J. Jiang, "Impact of antenna height on propagation characteristics of 2.4 ghz wireless channel in wheat fields," *Transactions of the Chinese Society of Agricultural Engineering*, vol. 25, no. s2, 2009. (In Chinese with English abstract).
- [59] S. Li and H. Gao, "Propagation characteristics of 2.4ghz wireless channel in cornfields," in *2011 IEEE 13th International Conference on Communication Technology*, pp. 136–140, 2011.
- [60] J. Li and F. Wang, "Modeling for wireless channel transmission loss characteristics under rapeseed growth big field environment," *Application Research of Computers*, vol. 34, no. 4, 2017. (In Chinese with English abstract).
- [61] Z. Gao, W. Li, Y. Zhu, Y. Tian, F. Pang, W. Cao, and J. Ni, "Wireless Channel Propagation Characteristics and Modeling Research in Rice Field Sensor Networks," *Sensors*, vol. 18, no. 9, 2018.
- [62] X. Ji, N. Wan, and J. Jiang, "Ecological parameters and effective distance of insecticidal lights," *Chinese Journal of Applied Entomology*, vol. 48, no. 3, pp. 669–674, 2011. (In Chinese with English abstract).
- [63] X. Jin, X. Tan, J. Liu, and J. Tang, "Application of frequency trembler pest-killing lamp in the major crops," *Guizhou Agricultural Sciences*, vol. 37, no. 2, pp. 53–54, 2009. (In Chinese with English abstract).
- [64] L. Bian, X. Cai, and Z. Chen, "Effects of light-emitting diode light traps with a suction fan on the population dynamics of the tea leafhopper *Empoasca onukii* (hemiptera: Cicadellidae) within the effective distance," *Journal of Plant Protection*, vol. 46, no. 4, pp. 902–909, 2019. (In Chinese with English abstract).
- [65] C. Feng and L. Liu, "Preliminary study on solar double-wave trapping light in rice field," *China Rice*, vol. 20, no. 6, pp. 62–65, 2014. (In Chinese with English abstract).
- [66] Proakis, *Digital Communications 5th Edition*. McGraw Hill, 2007.
- [67] M. K. Simon and M. Alouini, *Digital Communication over Fading Channels 2th Edition*. John Wiley & Sons, 2005.
- [68] A. F. Molisch, *Wireless Communications*. Wiley-IEEE Press, 2005.
- [69] P. He, S. Ying, W. Xin, and T. Li, "Modeling wireless sensor networks radio frequency signal loss in corn environment," *Multimedia Tools & Applications*, vol. 76, no. 19, pp. 19479–19490, 2016.
- [70] D. Balachander, T. R. Rao, and G. Mahesh, "Rf propagation investigations in agricultural fields and gardens for wireless sensor communications," in *Information & Communication Technologies (ICT), 2013 IEEE Conference on*, 2013.
- [71] S. Pirzada, *An Introduction to Graph Theory*. Jun 2012.
- [72] T. Shuai and X. Hu, "Connected set cover problem and its applications," in *Algorithmic Aspects in Information and Management*, (Berlin, Heidelberg), pp. 243–254, Springer Berlin Heidelberg, 06 2006.
- [73] W. Zhang, W. Wu, W. Lee, and D.-Z. Du, "Complexity and approximation of the connected set-cover problem," *Journal of Global Optimization*, vol. 53, no. 3, pp. 563–572, 2012.
- [74] X. Song, C. Cheng, and C. Zhou, "An analysis and investigation of algorithms for identifying convexity-concavity of a simple polygon," *Remote Sensing for Land & Resources*, vol. 19, no. 3, pp. 25–31, 2011. (In Chinese with English abstract).
- [75] S. Kashi, "Area coverage of heterogeneous wireless sensor networks in support of internet of things demands," *Computing*, vol. 101, 05 2018.
- [76] L. Shu, M. Mukherjee, and X. Wu, "Toxic gas boundary area detection in large-scale petrochemical plants with industrial wireless sensor networks," *IEEE Communications Magazine*, vol. 54, no. 10, pp. 22–28, 2016.
- [77] R. Liu, "A simple and fast algorithm for detecting the convexity and concavity of vertices for an arbitrary polygon," *Journal of Software*, vol. 13, no. 7, pp. 1309–1312, 2002. (In Chinese with English abstract).
- [78] L. Erhan, M. Nduvaku, M. Di Mauro, W. Song, M. Chen, G. Fortino, O. Bagdasar, and A. Liotta, "Smart anomaly detection in sensor systems: A multi-perspective review," *Information Fusion*, vol. 67, pp. 64–79, 2021.
- [79] F. Cauteruccio, G. Fortino, A. Guerrieri, A. Liotta, D. C. Mocanu, C. Perra, G. Terracina, and M. Torres Vega, "Short-long term anomaly detection in wireless sensor networks based on machine learning and multi-parameterized edit distance," *Information Fusion*, vol. 52, pp. 13–30, 2019.
- [80] F. Aiello, G. Fortino, A. Guerrieri, and R. Gravina, "Maps: A mobile agent platform for wsns based on java sun spots," *Third International Workshop on Agent Technology for Sensor Networks*, pp. 41–48, 2009.
- [81] X. Wang, S. Wang, and A. Jiang, "Optimized deployment strategy of mobile agents in wireless sensor networks," in *Sixth International Conference on Intelligent Systems Design and Applications*, vol. 2, pp. 893–898, 2006.
- [82] M. Pugliese, L. Pomante, and F. Santucci, "Topology optimization and network deployment algorithm in wsns for mobile agent-based applications," in *2010 Fourth UKSim European Symposium on Computer Modeling and Simulation*, pp. 460–465, 2010.



**Fan Yang** received the B.S. degree in communication engineering from Nanjing University of Posts and Telecommunications, Nanjing, China, in 2011 and the M.S. degree in computer application technology from Nanjing Agricultural University, Nanjing, China, in 2014. He is currently working toward the Ph.D. degree with the College of Engineering, Nanjing Agricultural University, China. His research interests include performance analysis and protocol design on networks, Wireless Sensor Networks and IoT.



**Lei Shu** [M'07- SM'15] received the B.S. degree in computer science from South Central University for Nationalities, China, in 2002, and the M.S. degree in computer engineering from Kyung Hee University, South Korea, in 2005, and the Ph.D. degree from the Digital Enterprise Research Institute, National University of Ireland, Galway, Ireland, in 2010. Until 2012, he was a Specially Assigned Researcher with the Department of Multimedia Engineering, Graduate School of Information Science and Technology, Osaka University, Japan. He is currently a Distinguished Professor with the Nanjing Agricultural University, China, and a Lincoln Professor with the University of Lincoln, U.K. He is also the Director of the NAU-Lincoln Joint Research Center of Intelligent Engineering. He has published over 400 papers in related conferences, journals, and books in the areas of sensor networks and Internet of Things. His current H-index is 60 and i10-index is 235 in Google Scholar Citation. His current research interests include wireless sensor networks and Internet of Things. He has also served as a TPC member for more than 150 conferences, DCOSS, MASS, ICC, GLOBECOM, ICCCN, WCNC, and ISCC. He has served over 50 various Co-Chair for international conferences/workshops, such as IWCMC, ICC, ISCC, ICNC, Chinacom, especially the Symposium Co-Chair for IWCMC 2012, ICC 2012, the General Co-Chair for Chinacom 2014, Qshine 2015, Collaboratecom 2017, DependSys 2018, and SCI 2019, the TPC Chair for InisCom 2015, NCCA 2015, WICON 2016, NCCA 2016, Chinacom 2017, InisCom 2017, WMNC 2017, and NCCA 2018. He has also served on the Editorial Boards, including the IEEE Transactions on Industrial Informatics, IEEE Communications Magazine, IEEE Network Magazine, IEEE Systems Journal, IEEE ACCESS, IEEE/CJA JAS.



**Yuli Yang** [S'04-M'08-SM'19] received her Ph.D. degree in Communications & Information Systems from Peking University in July 2007. Since Dec 2019, she has been with the University of Lincoln as a Senior Lecturer in Electrical/Electronic Engineering. From Jan 2010 to Dec 2019, she was with King Abdullah University of Science & Technology, Melikşah University, and the University of Chester on various academic positions. Her industry experience includes working as a Research Scientist with Bell Labs Shanghai, from Aug 2007 to Dec 2009, and an Intern Researcher with Huawei Technologies, from June 2006 to July 2007. Her research interests include modelling, design, analysis and optimization of wireless systems and networks.



**Guangjie Han** [S'03-M'05-SM'18] received the Ph.D. degree from Northeastern University, Shenyang, China, in 2004. He is currently a Professor with the College of Engineering, Nanjing Agricultural University, Nanjing, China. He is also a Professor with the Department of Information and Communication System, Hohai University, Changzhou, China, and a Professor with the State Key Laboratory of Acoustics, Institute of Acoustics, Chinese Academy of Sciences, Beijing, China. He has authored more than 330 papers published

in related international conference proceedings and journals, including the IEEE COMST, IEEE TII, IEEE TMC, IEEE TVT, IEEE TIE, IEEE TPDS, IEEE TETC, IEEE IoT Journal, IEEE TETCI, IEEE TCC, IEEE Systems, IEEE Sensors, IEEE Wireless Communications, IEEE Communications, IEEE Network. He is the holder of 130 patents. His current H-index is 45 and i10-index is 165 in Google Citation (Google Scholar). Total citation of his papers by other people is more than 5200 times. His current research interests include Internet of Things, the industrial Internet, mobile computing, artificial intelligence, and security. He was a recipient of ComManTel 2014, ComComAP 2014, Chinacom 2014, and Qshine 2016 Best Paper Awards. He has served as a Co-Chair for more than 50 conferences/workshops and as a Technical Program Committee Member of more than 150 conferences. He has served on the Editorial Boards of up to 16 international journals, including the IEEE JSAC, IEEE Network, IEEE Systems, IEEE ACCESS, IEEE/CCA JAS, Telecommunication Systems. He has guest edited a number of special issues in IEEE Journals and Magazines, including the IEEE Communications, IEEE Wireless Communications, IEEE Transactions on Industrial Informatics, Computer Networks. He has served as a Reviewer of more than 60 journals.



**Simon Pearson** is currently a Professor and the Director of the Lincoln Institute of Agri-Food Technology (LIAT) in the College of Science at the University of Lincoln, U.K. He was the CEO of a large UK farming company, worked for Marks and Spencer plc in a technical role and as a scientist at the University of Reading. His main research interests include a diverse range of agricultural technology applications, consisting of robotic systems, automation, energy control and management, food safety systems and novel crop development.



**Kailiang Li** received the B.S. degrees from Guangdong University of Petrochemical Technology, Maoming, China, in 2015. He is currently a Research Assistant at College of Engineering, Nanjing Agricultural University. His research interests are Wireless Sensor Networks and Internet of Things. He has published 11 papers in related conferences, journals in the areas of sensor networks and Internet of Things.

## APPENDIX A PROOF OF LEMMA 1

Lemma 1 can be proved by contradiction. Suppose  $|J(V)| \leq 2$  and assume the network has more than one coverage holes. Let  $H = \{H_1, H_2, \dots, H_n\}$  be a set of coverage holes where  $n \geq 2$ . From the *Definition 13*, each coverage hole have two BIs and any two coverage holes have at most one common BIs, so the lower bound on BIs is  $(n+1)$ . According to *Definition 11*, BI is the BP whose index in  $B$  is equal to JP. Thus, the number of BIs is the same as that of JPs, i.e.,  $|J(V)| = n+1$ . Due to  $n \geq 2$ , we have  $|J(V)| \geq 3$ , which is contradictory with the initial assumption that  $|J(V)| \leq 2$ .

## APPENDIX B PROOF OF LEMMA 2

In this appendix, we derive the expression for the maximum transmission range  $R_c^{s_i}$  of the SIL node  $s_i$ .

As is known, if  $s_j$  can successfully receive the signals sent from  $s_i$ , the average received SNR at  $s_j$  must exceed the receiver's sensitivity threshold  $\gamma_{th}$ , i.e.,

$$\bar{\phi}_{i,j} = \frac{P_i \mathbb{E}(\theta_{i,j}^2)}{\sigma^2} = \frac{P_i}{\sigma^2} \left( \frac{d_0}{d(s_i, s_j)} \right)^{\alpha_i} \geq \gamma_{th}. \quad (\text{A.1})$$

Thus, we have

$$d(s_i, s_j) \leq d_0 \sqrt[\alpha_i]{\frac{P_i}{\sigma^2 \gamma_{th}}}. \quad (\text{A.2})$$

Let  $R_c^{s_i}$  denote the maximum transmission distance of SIL node  $s_i$ , within which  $s_j$  is able to successfully receive the signals sent from  $s_i$ . Therefore,  $R_c^{s_i}$  can be expressed as

$$R_c^{s_i} = d_0 \sqrt[\alpha_i]{\frac{P_i}{\sigma^2 \gamma_{th}}}. \quad (\text{A.3})$$